ims

سیستم مدیریت ایزو

www.isomanagement.ir

تماس تلفنی جهت دریافت مشاوره:

۱. مشاور دفتر تهران (آقای محسن ممیز)

۰۹۱۲ ۹۶۳ ۹۳۳۶

۲. مشاور دفتر اصفهان (سرکار خانم لیلا ممیز)

۰۹۱۳ ۳۲۲ ۸۲۵۹

مجموعه **سیستم مدیریت ایزو** با هدف بهبود مستمرعملکرد خود و افزایش رضایت مشتریان سعی بر آن داشته، کلیه استانداردهای ملی و بین المللی را در فضای مجازی نشر داده و اطلاع رسانی کند، که تمام مردم ایران از حقوق اولیه شهروندی خود آگاهی لازم را کسب نمایند و از طرف دیگر کلیه مراکز و کارخانه جات بتوانند به راحتی به استانداردهای مورد نیاز دسترسی داشته باشند.

این موسسه اعلام می دارد در کلیه گرایشهای سیستم های بین المللی ISO پیشگام بوده و کلیه مشاوره های ایزو به صورت رایگان و صدور گواهینامه ها تحت اعتبارات بین المللی سازمان جهانی IAF و تامین صلاحیت ایران می باشد.

هم اکنون سیستم خود را با معیارهای جهانی سازگار کنید...

ISO

9001
14001
18001

CE

GMP
GAP
22716

IMS
HSE-MS

22000
HACCP

45001

INSO
15675
1st Edition
2019
Identical with
ISO/TR13587:2012

جمهوری اسلامی ایران
Islamic Republic of Iran
سازمان ملی استاندارد ایران
Iranian National Standardization Organization

استاندارد ملی ایران
۱۵۶۷۵
چاپ اول
۱۳۹۷

سه رویکرد آماری برای ارزیابی و تفسیر عدم‌قطعیت اندازه‌گیری

Three statistical approaches for the assessment and interpretation of measurement uncertainty

ICS: 03.120.30

استاندارد ملی ایران شمارهٔ ۱۵۶۷۵ (چاپ اول) : سال ۱۳۹۷

سازمان ملی استاندارد ایران

تهران، ضلع جنوب غربی میدان ونک، خیابان ولیعصر، پلاک ۲۵۹۲

صندوق‌پستی: ۶۱۳۹-۱۴۱۵۵ تهران- ایران

تلفن: ۵-۸۸۸۷۹۴۶۱

دورنگار: ۸۸۸۸۷۰۸۰ و ۸۸۸۸۷۱۰۳

کرج - شهر صنعتی، میدان استاندارد

صندوق پستی: ۱۶۳-۳۱۵۸۵ کرج - ایران

تلفن: ۸- ۳۲۸۰۶۰۳۱ (۰۲۶)

دورنگار: ۳۲۸۰۸۱۱۴ (۰۲۶)

رایانامه: standard@isiri.gov.ir

وبگاه: http://www.isiri.gov.ir

Iranian National Standardization Organization (INSO)

No. 2592Valiasr Ave., South western corner of Vanak Sq., Tehran, Iran

P. O. Box: 14155-6139, Tehran, Iran

Tel: + 98 (21) 88879461-5

Fax: + 98 (21) 88887080, 88887103

Standard Square, Karaj, Iran

P.O. Box: 31585-163, Karaj, Iran

Tel: + 98 (26) 32806031-8

Fax: + 98 (26) 32808114

Email: standard@isiri.gov.ir

Website: http://www.isiri.gov.ir

به نام خدا

آشنایی با سازمان ملی استاندارد ایران

سازمان ملی استاندارد ایران به موجب بند یک مادهٔ ۳ قانون اصلاح قوانین و مقررات مؤسسهٔ استاندارد و تحقیقات صنعتی ایران، مصوب بهمن ماه ۱۳۷۱ تنها مرجع رسمی کشور است که وظیفهٔ تعیین، تدوین و نشر استانداردهای ملی (رسمی) ایران را به عهده دارد.

تدوین استاندارد در حوزه‌های مختلف در کمیسیون‌های فنی مرکب از کارشناسان سازمان، صاحب‌نظران مراکز و مؤسسات علمی، پژوهشی، تولیدی و اقتصادی آگاه و مرتبط انجام می‌شود و کوششی همگام با مصالح ملی و با توجه به شرایط تولیدی، فناوری و تجاری است که از مشارکت آگاهانه و منصفانهٔ صاحبان حق و نفع، شامل تولیدکنندگان، مصرف‌کنندگان، صادرکنندگان و واردکنندگان، مراکز علمی و تخصصی، نهادها، سازمان‌های دولتی و غیردولتی حاصل می‌شود. پیش‌نویس استانداردهای ملی ایران برای نظرخواهی به مراجع ذی‌نفع و اعضای کمیسیون‌های مربوط ارسال می‌شود و پس از دریافت نظرها و پیشنهادها در کمیتهٔ ملی مرتبط با آن رشته طرح و در صورت تصویب، به عنوان استاندارد ملی (رسمی) ایران چاپ و منتشر می‌شود.

پیش‌نویس استانداردهایی که مؤسسات و سازمان‌های علاقه‌مند و ذی‌صلاح نیز با رعایت ضوابط تعیین شده تهیه می‌کنند درکمیته ملی طرح، بررسی و درصورت تصویب، به عنوان استاندارد ملی ایران چاپ و منتشر می‌شود. بدین ترتیب، استانداردهایی ملی تلقی می‌شود که بر اساس مقررات استاندارد ملی ایران شمارهٔ ۵ تدوین و در کمیتهٔ ملی استاندارد مربوط که در سازمان ملی استاندارد ایران تشکیل می‌شود به تصویب رسیده باشد.

سازمان ملی استاندارد ایران از اعضای اصلی سازمان بین‌المللی استاندارد (ISO)[1]، کمیسیون بین‌المللی الکتروتکنیک (IEC)[2] و سازمان بین‌المللی اندازه‌شناسی قانونی (OIML)[3] است و به عنوان تنها رابط[4] کمیسیون کدکس غذایی (CAC)[5] در کشور فعالیت می‌کند. در تدوین استانداردهای ملی ایران ضمن توجه به شرایط کلی و نیازمندی‌های خاص کشور، از آخرین پیشرفت‌های علمی، فنی و صنعتی جهان و استانداردهای بین‌المللی بهره‌گیری می‌شود.

سازمان ملی استاندارد ایران می‌تواند با رعایت موازین پیش‌بینی شده در قانون، برای حمایت از مصرف‌کنندگان، حفظ سلامت و ایمنی فردی و عمومی، حصول اطمینان از کیفیت محصولات و ملاحظات زیست‌محیطی و اقتصادی، اجرای بعضی از استانداردهای ملی ایران را برای محصولات تولیدی داخل کشور و/یا اقلام وارداتی، با تصویب شورای عالی استاندارد، اجباری کند. سازمان می‌تواند به منظور حفظ بازارهای بین‌المللی برای محصولات کشور، اجرای استاندارد کالاهای صادراتی و درجه‌بندی آن را اجباری کند. همچنین برای اطمینان بخشیدن به استفاده‌کنندگان از خدمات سازمان‌ها و مؤسسات فعال در زمینهٔ مشاوره، آموزش، بازرسی، ممیزی و صدورگواهی سیستم‌های مدیریت کیفیت و مدیریت زیست‌محیطی، آزمایشگاه‌ها و مراکز واسنجی (کالیبراسیون) وسایل سنجش، سازمان ملی استاندارد این‌گونه سازمان‌ها و مؤسسات را بر اساس ضوابط نظام تأیید صلاحیت ایران ارزیابی می‌کند و در صورت احراز شرایط لازم، گواهینامهٔ تأیید صلاحیت به آن‌ها اعطا و بر عملکرد آن‌ها نظارت می‌کند. ترویج دستگاه بین‌المللی یکاها، واسنجی وسایل سنجش، تعیین عیار فلزات گرانبها و انجام تحقیقات کاربردی برای ارتقای سطح استانداردهای ملی ایران از دیگر وظایف این سازمان است.

1- International Organization for Standardization
2- International Electrotechnical Commission
3- International Organization for Legal Metrology (Organisation Internationale de Metrologie Legals)
4- Contact point
5- Codex Alimentarius Commission

کمیسیون فنی تدوین استاندارد

«سه رویکرد آماری برای ارزیابی و تفسیر عدم‌قطعیت اندازه گیری»

رئیس:	**سمت و/یا محل اشتغال:**
مرادی، محسن (کارشناسی ارشد فیزیک)	اداره کل استاندارد آذربایجان غربی
دبیر:	
علیزاده، محمدرضا (کارشناسی ارشد مدیریت صنعتی)	اداره کل استاندارد استان آذربایجان غربی
اعضا: (اسامی به ترتیب حروف الفبا)	
ادیبی، جواد (کارشناسی ارشد مهندسی نرم‌افزار)	اداره کل استاندارد استان آذربایجان غربی
اسگویی، الناز (دکتری ریاضی)	دانشگاه صنعتی ارومیه
اوحدی، افشین (کارشناسی ارشد مهندسی کشاورزی)	سازمان ملی استاندارد ایران
حسین‌زاده، حسین (کارشناسی ارشد مدیریت صنعتی)	عضو مستقل
دباغ، رحیم (دکتری اقتصاد)	دانشگاه صنعتی ارومیه
رنجبریان، رسول (دکتری مدیریت دولتی)	دانشگاه آزاداسلامی- واحد شاهین‌دژ
هوشیار قهرمانلو، خیراله (دکتری آمار کاربردی)	دانشگاه ارومیه
هوشیار قهرمانلو، نوشین (دکتری ریاضی)	دانشگاه غیرانتفاعی کمال ارومیه

اعضا: (اسامی به ترتیب حروف الفبا)	**سمت و/یا محل اشتغال:**
نیازی، علیرضا (کارشناسی ارشد مدیریت اجرایی)	اداره کل استاندارد استان آذربایجان غربی
ویراستار:	
اوحدی، افشین (کارشناسی ارشد مهندسی کشاورزی)	سازمان ملی استاندارد ایران

فهرست مندرجات

عنوان **صفحه**

عنوان **صفحه**

پیش‌گفتار

استاندارد «سه رویکرد آماری برای ارزیابی و تفسیر عدم‌قطعیت اندازه‌گیری» که پیش‌نویس آن در کمیسیون‌های مربوط بر مبنای پذیرش استانداردهای بین‌المللی/منطقه‌ای به عنوان استاندارد ملی ایران به روش اشاره شده در مورد الف، بند۷ استاندارد ملی ایران، شمارۀ ۵ تهیه و تدوین شده، در دویست و بیست و هشتمین اجلاسیه کمیته ملی استاندارد مدیریت کیفیت مورخ ۹۷/۱۲/۲۰ تصویب شد. اینک این استاندارد به استناد بند یک مادۀ ۳ قانون اصلاح قوانین و مقررات استاندارد و تحقیقات صنعتی ایران، مصوب بهمن‌ماه ۱۳۷۱، به عنوان استاندارد ملی ایران منتشر می‌شود.

استانداردهای ملی ایران بر اساس استاندارد ملی ایران شمارۀ ۵ (استانداردهای ملی ایران- ساختار و شیوه نگارش) تدوین می‌شوند. برای حفظ همگامی و هماهنگی با تحولات و پیشرفت‌های ملی و جهانی در زمینه صنایع، علوم و خدمات، استانداردهای ملی ایران در صورت لزوم تجدیدنظر خواهند شد و هر پیشنهادی که برای اصلاح یا تکمیل این استانداردها ارائه شود، در هنگام تجدیدنظر در کمیسیون فنی مربوط مورد توجه قرار خواهد گرفت. بنابراین، باید همواره از آخرین تجدیدنظر استاندارد ملی ایران استفاده کرد.

این استاندارد ملی بر مبنای پذیرش استاندارد بین‌المللی زیر به روش«معادل یکسان» تهیه و تدوین شده و شامل ترجمه تخصصی کامل متن آن به زبان فارسی می‌باشد و معادل یکسان استاندارد بین‌المللی مزبور است:

ISO/TR13587:2012, Three statistical approaches for the assessment and interpretation of measurement uncertainty

مقدمه

پذیرش استاندارد (GUM) ISO/IEC 98-3 [۱] منجر به افزایش شناخت نیازهای مربوط به مفاهیم عدم‌قطعیت در نتایج اندازه‌گیری شده است. تأیید صلاحیت آزمایشگاه‌ها بر اساس استانداردهای بین‌المللی نظیر17025 ISO [۲] این فرآیند را سرعت می‌بخشد. .تشخیص این‌که مفهوم عدم‌قطعیت در تصمیم‌گیری کارآمد موردنیاز است، اندازه‌شناسان در تمام انواع آزمایشگاه‌ها، از مؤسسات ملی اندازه‌شناسی گرفته تا آزمایشگاه‌های کالیبراسیون تجاری را به انجام تلاش‌هایی برای توسعه ارزیابی‌های عدم‌قطعیت مناسب برای انواع مختلف اندازه‌گیری‌ها با استفاده از روش‌های گفته شده در استاندارد GUM (راهنمای عدم‌قطعیت در اندازه‌گیری)[1] وا داشته است.

بعضی از نقاط قوت مربوط به رویه‌های که در استاندارد GUM مشخص شده و مورد استقبال واقع شده‌اند، عبارت هستند از: رویکرد استاندارد شده برای ارزشیابی عدم‌قطعیت، سازگار بودن منابع عدم‌قطعیت که هم از لحاظ آماری (نوع A) یا غیر آماری (نوعB) ارزشیابی شده‌اند و تأکید آن‌ها بر گزارش همه منابع عدم‌قطعیت مربوط به آن‌ها. رویکرد اصلی در نشر مفهوم عدم‌قطعیت در استاندارد GUM مبتنی بر تقریب خطی تابع اندازه‌گیری است که عموماً به‌راحتی به‌دست می‌آید و در بسیاری ازموقعیت‌های عملی نتایجی به دست می‌دهد که شبیه نتایجی است که به‌طور رسمی به‌دست آمده‌اند. به‌طور خلاصه از زمان به کارگیری استاندارد GUM، انقلابی در ارزشیابی عدم‌قطعیت به‌وجود آمده است.

البته همیشه نیاز به تلاش‌های بیشتری جهت تقویت ازرشیابی عدم‌قطعیت در موقعیت‌های کاربردی خاص و توسعه نواحی تحت پوشش وجود دارد. از جمله کارهای دیگر، تشکیل کمیته مشترک برای رهنمودها در اندازه‌شناسی(JCGM)[2] که مسئولیت استاندارد GUM را از سال ۲۰۰۰ بر عهده داشته، بوده است، که پیوست ۱ را دراستاندارد GUM کامل کرده است برای مثال « نشر توزیع‌ها با استفاده از روش مونت‌کارلو[3] » (به عنوان GUMS1 نامیده می‌شود) [۳]. JCGM در حال توسعه پیوست‌های دیگری به استاندارد GUM در موضوعاتی نظیر مدل‌سازی و مدل‌ها با سایر کمیت‌های خروجی است.

به‌دلیل آن که طیف وسیعی از مشکلات احتمالی اندازه‌گیری ممکن در تعریف عدم‌قطعیت اندازه‌گیری در استاندارد ISO/IEC Guide 99:2007 [۴] به‌عنوان پارامتر مشخص‌کننده غیرمنفی در پراکندگی ارزش‌های مقداری نسبت به مقدار اندازه‌ده که بر اساس اطلاعات استفاده شده است نمی‌تواند به‌طور منطقی در بیش از یک سطح مفهومی به‌کار گرفته شود، در نتیجه، تعریف و درک نقش‌های مناسب کمیت‌های آماری مختلف که در ارزشیابی عدم‌قطعیت حتی برای کاربردهای اندازه‌گیری بدیهی هستند، نیز موضوعی است که هم به‌طور نسبی در زمینه اندازه‌شناسی و هم در زمینه آماری نیازمند توجه ویژه است.

1 - Guide Uncertainty Measurement

2 - Joint Committee for Guides in Metrology

3 - Monte Carlo

تحقیقات اخیر این روش را از دیدگاه اندازه‌شناسی دنبال کرده‌اند، برخی از نویسندگان نیز بر خصوصیات روش‌های اجرایی آماری ارائه شده دراستاندارد GUM متمرکز شده‌اند. منبع [۵] نشان می‌دهد که این روش با هریک از تفسیرهای روش‌های بیزی[1] یا فراوانی‌گرا، موکداً سازگار نیستند. منبع [۶] برخی اصلاحات جزئی در روش استاندارد GUM را که نتایج را با تفسیرهای بیزی سازگار می‌کند مورد بررسی قرار می‌دهد. منبع [۷] رابطه بین روش‌های ارزشیابی عدم‌قطعیت پیشنهاد شده در GUMS1 و نتایج تحلیل بیزی را برای یک رده خاص از مدل‌های ویژه بحث می‌کند. منبع [۸] هم‌چنین تفسیرهای احتمالات مختلفی از بازه تحت پوشش و تقریب پیشنهادی توزیعی پسین را برای این رده از تحلیل‌های بیزی به‌وسیله توزیع‌های احتمال از خانواده توزیع پیرسون[2] را بحث می‌کند.

منبع [۹] رویکردهای فراوانی‌گرا (متعارف) و بیزی را برای ارزشیابی عدم‌قطعیت فراهم می‌کند. با این حال، مطالعه در نظام‌های اندازه‌گیری برای تمامی منابع عدم‌قطعیت محدود است و می‌تواند با استفاده از روش‌های نوع A مورد ارزشیابی قرار گیرد. در طرف مقابل، نظام‌های اندازه‌گیری با منابع عدم‌قطعیت با استفاده از روش‌های نوع A و B ارزشیابی می‌شوند که در این استاندارد مورد بررسی قرار گرفته و با استفاده از چندین مثال، از جمله مثال‌های پیوست H از استاندارد GUM نشان داده شده است.

متخصصان علم آمار از لحاظ پیشینه، تأکید زیادی بر روی استفاده از روش‌هایی برای ارزشیابی عدم‌قطعیت که قابلیت قضاوت یا تفسیر احتمالی داشته باشند،کرده‌اند. با استفاده از کارهای ایشان، اغلب اندازه‌شناسی بیرونی، رویکردهای متفاوت زیادی برای استنباط آماری مرتبط با ارزشیابی عدم‌قطعیت توسعه یافته است. این استاندارد برخی از این رویکردها را برای ارزشیابی عدم‌قطعیت از منظر آماری و ارتباط آن‌ها را با مدل‌هایی که در حال حاضر در اندازه‌شناسی مورد استفاده قرار می‌گیرد یا این که درجامعه اندازه‌شناسی در حال توسعه هستند ارائه می‌دهد. رویکردهای آماری خاص که از طریق آن‌ها روش‌های متفاوتی برای ارزشیابی عدم‌قطعیت توضیح داده می‌شود شامل رویکردهای فراوانی‌گرا، بیزی و اتکایی[3] است که بعد از مشخص شدن کنوانسیون‌های نمادگذاری مورد نیاز برای تعیین انواع مختلف کمیت‌ها به‌طور مفصل مورد بحث و بررسی قرار می‌گیرند.

1 - Bayesian
2- Pearson
3- Fiducial

سه رویکرد آماری برای ارزیابی و تفسیر عدم‌قطعیت اندازه‌گیری

۱ هدف و دامنهٔ کاربرد

هدف از تدوین این استاندارد، تعیین سه رویکرد آماری اساسی برای ارزشیابی و تفسیر عدم‌قطعیت اندازه‌گیری است. رویکرد فراوانی‌گرا شامل بازه‌های عدم‌قطعیت خودگردان[1]، رویکرد بیزی و رویکرد استنباط اتکایی است. ویژگی متداول این رویکردها در تعیین شفاف و واضح تفسیرهای احتمالی یا تنظیم نتایج بازه عدم‌قطعیت کاربرد دارد. برای هر رویکرد، روش اساسی توضیح داده شده و مفروضات بنیادی و اساسی و تفسیر احتمالی نتیجه عدم‌قطعیت مورد بحث قرار می‌گیرد. هر یک از این رویکردها با استفاده از دو مثال مورد کاربرد در استانداردهای ISO/IEC Guide98-3 (عدم‌قطعیت در اندازه‌گیری- قسمت ۳: راهنما برای حالت عدم‌قطعیت در اندازه‌گیری (GUM:1995)) نشان داده می‌شوند. همچنین، این استاندارد شامل یک بحث در ارتباط بین روش‌های پیشنهادی در پیوست ۱ استاندارد GUM و این سه رویکرد آماری است.

۲ مراجع الزامی

در مراجع زیر ضوابطی وجود دارد که در متن این استاندارد به صورت الزامی به آن‌ها ارجاع داده شده است. بدین‌ترتیب، آن ضوابط جزئی از این استاندارد محسوب می‌شوند.

در صورتی که به مرجعی با ذکر تاریخ انتشار ارجاع داده شده باشد، اصلاحیه‌ها و تجدیدنظرهای بعدی آن برای این استاندارد الزام‌آور نیست. در مورد مراجعی که بدون ذکر تاریخ انتشار به آن‌ها ارجاع داده شده است، همواره آخرین تجدیدنظر و اصلاحیه‌های بعدی برای این استاندارد الزام‌آور است.

استفاده از مراجع زیر برای کاربرد این استاندارد الزامی است:

2-1 ISO 3534-1, Statistics — Vocabulary and symbols — Part 1: General statistical terms and terms used in probability

2-2 ISO 3534-2, Statistics — Vocabulary and symbols — Part 2: Applied statistics

2-3 ISO/IEC Guide 98-3, Uncertainty of measurement — Part 3: Guide to the expression of uncertainty in measurement (GUM:1995)

2-4 ISO/IEC Guide 98-3/Suppl 1:, Uncertainty of measurement — Part 3: Guide to the expression of uncertainty in measurement (GUM:1995) — Supplement 1: Propagation of distributions using a Monte Carlo

1 - Bootstrap

۳ اصطلاحات و تعاریف

در این استاندارد، علاوه بر اصطلاحات و تعاریف بـه کـار رفتـه در اسـتانداردهای ISO3534-1,ISO3534-2 موارد زیر نیز کاربرد دارد.

۱-۳

تابع توزیع تجربی **empirical distribution function**

تابع توزیع تجمعی تجربی **empirical cumulative distribution function**

تابع توزیعی که احتمال $1/n$ را برای هر یک از اقلام n در یک الگوی تصادفی اختصاص می‌دهد. برای مثـال، تابع توزیع تجربی یک تابع یک مرحله‌ای است که به صورت زیر تعریف می‌شود:

$$F_n(x) = \frac{|\{x_i \le x\}|}{n},$$

که در آن:

$\{x_1,\ldots,x_n\}$ نمونه و $|A|$ تعداد مؤلفه‌ها در مجموعه A می‌باشد.

۲-۳

تحلیل حساسیت بیزی **Bayesian sensitivity analysis**

مطالعه اثر انتخاب توزیع پیشین برای مدل پارامترهای آماری در توزیع پسین اندازه‌ده است

۳-۳

آمار بسنده **sufficient statistic**

تابعی از یک نمونه تصادفی $X_1,\ldots,X_n$ از یک تابع چگالی احتمال با پارامترθ برای توزیـع شـرطی $X_1,\ldots,X_n$ که این تابع به θ وابسته نیست.

یادآوری- یک آماره بسنده حاوی مقادیر زیادی اطلاعات در باره θ از $X_1,\ldots,X_n$ است.

۴-۳

مدل مشاهده **observation model**

رابطه ریاضی بین مجموعه‌ای از اندازه‌گیری‌ها (علائم)، اندازه‌ده و خطاهای اندازه‌گیـری تصـادفی مـرتبط بـا آن‌ها است.

۵-۳

معادله ساختاری **structural equation**

مدل آماری مربوط به متغیر تصادفی قابل مشاهده با پارامترهای نامعلوم و متغیر تصادفی غیرقابل مشاهده که توزیع آن معلوم بوده و مستقل از پارامترهای نامعلوم است.

۶-۳

توزیع کای‌دوی[۱] غیر مرکزی **non-central chi-squared distribution**

توزیع احتمالی که کای‌دوی نوعی (یا مرکزی) را ایجاد می‌کند.

یادآوری۱- برای K مستقل، متغیر تصادفی توزیع نرمال X_i با میانگین μ_i و واریانس σ_i^2 متغیر تصادفی $X^2=\sum_{i=1}^{k}(X_i/\sigma_i)^2$ توزیع کای‌دو غیرمرکزی است. توزیع کای‌دو غیرمرکزی دو پارامتر دارد: k، درجه آزادی (به‌عنوان مثال، تعداد X_i) و λ، که مرتبط با میانگین متغیرهای تصادفی X_i با $\lambda=\sum_{i=1}^{k}(\mu_i/\sigma_i)^2$ که پارامتر غیرمرکزی نامیده می‌شود.

یادآوری۲- تابع چگالی احتمال متناظر به عنوان ترکیب تابع چگالی احتمال χ^2 مرکزی به‌صورت زیر است:

$$g_x(\xi)=\sum_{i=0}^{\infty}\frac{e^{-\lambda/2}(\lambda/2)^i}{i!}g_{Y_{k+2i}}(\xi)$$
$$=\frac{e^{\frac{-(\xi+\lambda)}{2}}}{2^{\frac{k}{2}}}\sum_{i=0}^{\infty}\frac{\xi^{\frac{k}{2}+i-1}\lambda^i}{\Gamma\left(\frac{k}{2}+i\right)2^{2i}i!}$$

که Yq به‌عنوان کای‌دو با q درجه آزادی توزیع می‌شود.

1- Chi-squared

۴ نمادها (و اصطلاحات کوته نوشته)

در زیربند ۴-۱-۱ از استاندارد GUM، بیان شده که حروف لاتین هم برای نشان دادن کمیت فیزیکی که به‌وسیله اندازه‌گیری تعیین می‌شوند (به‌عنوان مثال: اندازه‌ده‌ها در استاندارد واژگانGUM) و هم‌چنین متغیرهای تصادفی که ممکن است مقادیر مشاهده‌ای مختلفی از کمیت فیزیکی را نشان دهند مورد استفاده قرار می‌گیرند. استفاده از نمادهای مشابه، که با معنای مختلفی فقط در متن مشخص شده‌اند، می‌توانند از لحاظ تفسیر مشکل به‌نظر برسد و گاهی اوقات نیز منجر به ابهام یا سوءتفاهم‌های غیرضروری می‌شود. برای از بین بردن این منبع بالقوه سردرگمی، نمادگذاری مرسوم که اغلب در ادبیات آماری استفاده می‌شود، در این استاندارد کاربرد دارد. در این استاندارد، حروف یونانی برای نشان دادن پارامترها در مدل آماری (به‌عنوان مثال، اندازه‌ده‌ها) دایر بر هم متغیرهای تصادفی و یا متغیر ثابت وابسته به رویکرد آماری در مدل مورد استفاده در ماهیت مدل هستند. حروف بزرگ لاتین برای نشان دادن متغیرهای تصادفی دایر بر مقادیر مختلفی از کمیت قابل مشاهده (به‌عنوان مثال، مقادیر قابل اندازه‌گیری بالقوه) و حروف کوچک لاتین برای نشان دادن مقادیر قابل مشاهده خاص در یک کمیت (به‌عنوان مثال، مقادیر اندازه‌گیری شده خاص) استفاده می‌شوند. از آنجایی که نمادگذاری‌های اضافی ممکن است نیازمند مشخص کردن سایر مفاهیم فیزیکی، ریاضی یا آماری باشند، هنوز هم برخی احتمالات برای ابهام وجود خواهد داشت.[1]؛ در این مواقع، متن تفسیر مناسب را ارائه خواهد داد.

۵ تبیین مسأله

۵-۱ در این استاندارد به بررسی مدل اندازه‌گیری که در آن: $\mu_1 \dots \mu_p$، کمیت‌های ورودی و θ کمیت خروجی است، پرداخته می شود:

$$\theta = f\left(\mu_1, \dots, \mu_p\right) \qquad (۱)$$

که در آن:

f به‌عنوان تابع اندازه‌گیری شناخته می‌شود.

تابع f به‌صورت روش ریاضی یا یک رویه محاسباتی مشخص می‌شود. در استاندارد GUM (یادآوری۱ زیربند ۴-۱) ارتباط تابعی مشابه به صورت زیر است:

۱- برای مثال تمام مقادیر با حروف یونانی در مدل آماری که باید پارامترها در مدل وجود داشته باشند، نشان داده نمی‌شوند. یک مثال رایج از این نوع کمیت، مجموعه‌ای از کمیت‌های غیرقابل مشاهده هستند، که نشان دهنده خطای اندازه‌گیری تصادفی است که در اکثر مدل‌های آماری پیدا می‌شوند (به‌عنوان مثال: ε_i در مدل $Y_i = \mu + \varepsilon_i$).

$$Y = f(X_1, \ldots, X_p) \tag{۲}$$

که نمی‌تواند به‌راحتی از اندازه‌گیری تابع ارزشیابی شده در مقادیر متغیرهای تصادفی متناظر برای هر ورودی مشاهده شده مشخص شود.

با استفاده از رویه توصیه شده در استاندارد GUM، P به‌عنوان کمیت نامعلوم $\mu_1 \ldots \mu_p$ هستند که به‌وسیله مقادیر $x_1, \ldots, x_p$ برآورد می‌شوند که با اندازه‌گیری فیزیکی یا از سایر منابع به‌دست می‌آیند. عدم‌قطعیت‌های استاندارد مرتبط با آن‌ها نیز از داده‌های مرتبط با روش‌های آماری یا توابع چگالی احتمال مبتنی بر دانش کارشناسان که متغیرها را مشخص می‌کنند، به‌دست می‌آید. هم‌چنین استاندارد GUM (به زیربند ۴-۵ منبع [۱۱] مراجعه شود) پیشنهاد می‌کند که مدل اندازه‌گیری مشابه که مرتبط با اندازه‌ده θ است به‌عنوان کمیت‌های ورودی $\mu_1, \ldots, \mu_p$ جهت محاسبه y از $x_1, \ldots x_p$ استفاده شود. بنابراین، مقدار اندازه‌گیری شده (یا در اصطلاحات آماری، برآورد) y از θ به صورت زیر به‌دست می‌آید:

$$y = f(x_1, \ldots, x_p), \tag{۳}$$

که Y ارزشیابی شده، $y = f(x_1, \ldots, x_p)$ برای اندازه‌گیری مقدار θ به‌کار می‌رود و برآورد مقادیر y، $x_1, \ldots, x_p$، به ترتیب با Y، $X_1, \ldots, X_p$ مرتبط هستند.

۵-۲ در این استاندارد، سه رویکرد آماری برای ارائه:

الف- بهترین برآورد y متعلق به θ؛

ب- عدم‌قطعیت استاندارد مرتبط $u(y)$ و

پ- یک بازه اطمینان یا بازه پوششی برای θ که برای تعیین یک احتمال پوششی (اغلب ٪۹۵) در نظر گرفته می‌شود، مورد استفاده قرار می‌گیرد.

۵-۳ هنگامی که از عدم‌قطعیت در استاندارد بحث می‌شود، بین عدم‌قطعیت‌های استاندارد ارزیابی شده مرتبط با برآوردهای کمیت‌های مختلف و مقادیر تئوری متناظر آن‌ها تمایز ایجاد می‌شود. بنابراین نمادی مانند σ_μ یا σ_x، عدم‌قطعیت استاندارد تئوری و نمادی مانند S_x یا s_x ، عدم‌قطعیت استاندارد ارزشیابی شده را قبل و بعد از مشاهده را به ترتیب نشان خواهند داد.

۶ رویکردهای آماری

۶-۱ رویکرد فراوانی‌گرا

۶-۱-۱ اولین رویکرد آماری که در آن عدم‌قطعیت را می‌توان از لحاظ احتمال، ارزشیابی کرد رویکرد فراوانی‌گرا است. رویکرد فراوانی‌گرا گاهی اوقات به‌عنوان رویکرد «کلاسیک » یا « متعارف» نیز نشان داده می‌شود. به هرحال با توجه به ماهیت عدم‌قطعیت در اندازه‌شناسی، این روش‌های شناخته شده باید اغلب برای به‌دست آوردن بازه عدم‌قطعیت فراوانی‌گرا تحت شرایط واقعی، سازگار باشد.

۶-۱-۲ در رویکرد فراوانی‌گرا، کمیت‌های ورودی $\mu_1,...,\mu_p$ در مدل اندازه‌گیری (۱) و کمیت خروجی θ به‌عنوان ثابت‌های نامعلوم در نظر گرفته می‌شوند. بنابراین داده‌های مربوط به هر پارامتر ورودی μ_i از برآورد مقدار ثابت θ مبتنی بر مدل‌های سنجش یا مدل‌های متناظر آماری به‌دست می‌آید و مورد استفاده قرار می‌گیرد. در خاتمه بازه اطمینان برای θ، برای یک سطح مشخصی از اطمینان با استفاده از یکی از اصول یا رویه‌های ریاضی چندگانه به‌دست می‌آید. برای مثال، حداقل مربعات، ماکسیمم درست‌نمایی یا خودگردان.

۶-۱-۳ از آنجایی که θ به‌عنوان یک ثابت عمل می‌کند حالت احتمالی مرتبط با بازه اطمینان برای θ یک حالت احتمال مستقیم در مورد مقدار آن نیست. در عوض، یک حالت احتمالی در مورد چگونگی تکرار رویه‌های استفاده شده برای به‌دست آوردن بازه عدم‌قطعیت برای اندازه‌ده‌ها است، که نتیجه آن استفاده تکراری مقادیر θ است. «استفاده تکراری» به این معنا است که ارزشیابی عدم‌قطعیت به دفعات مکرر با استفاده از داده‌های مختلف از توزیع مشابه گرفته شده است. بازه‌های عدم‌قطعیت فراوانی‌گرای مرسوم، یک حالت احتمالی در باره اجزای مراحل طولانی مدت رویه مورد استفاده برای تشکیل فاصله مجموعه‌ای از شرایط خاص فرض شده برای کاربرد فرآیند اندازه‌گیری است.

۶-۱-۴ در بسیاری از جنبه‌های عملی اندازه‌شناسی، از سوی دیگر، بازه‌های عدم‌قطعیت نشانگر عدم‌قطعیت مرتبط با برآوردهای به‌دست آمده با استفاده از کمیت‌های اندازه‌گیری شده (داده‌های مشاهده شده) است و هم‌چنین عدم‌قطعیت با برآورد کمیت‌های مبتنی بر دانش کارشناسان نیز ارتباط دارد. برای به‌دست آوردن بازه عدم‌قطعیت، همانند بازه‌های اطمینان، کمیت‌هایی که بر مبنای مقادیر سنجیده شده نیستند به‌عنوان متغیرهای تصادفی با توزیع احتمال برای مقادیرشان رفتار می‌کنند در حالی که مقادیری که ارزش‌های آن‌ها می‌تواند با استفاده از داده‌های آماری برآورد شوند، به‌عنوان یک نامعلوم ثابت مورد بررسی قرار می‌گیرند.

۶-۱-۵ در رویه‌های فراوانی‌گرای مرسوم برای ساختن بازه‌های اطمینان که بعداً برای به‌دست آوردن سطح اطمینان مشخصی تعدیل می‌شوند، بعد از میانگین‌گیری مقادیر بالقوه کمیت‌های ارزیابی شده با استفاده از قضاوت کارشناسان مشخص می‌شوند [۵]. چنین فواصل پوششی اصلاح شده حالت احتمالی طولانی مدت را درباره رویه‌های مورد استفاده برای به‌دست آوردن توزیع احتمالی بازه برای کمیت‌هایی که اندازه‌گیری نشده‌اند، همانند بازه اطمینان متداول که همه پارامترها را به صورت ثابت محسوب می‌کند، فراهم می‌کند.

۶-۱-۶ جدول ۱ مفاهیم و تفسیرهای رویکردهای فراوانی‌گرا، بیزی و اتکایی را برای ارزشیابی عدم‌قطعیت به‌طور خلاصه نشان می‌دهد.

جدول ۱- تفسیر رویکردهای ارزشیابی عدم‌قطعیت

رویکرد	تعیین ویژگی کمیت‌ها در مدل اندازه‌گیری	فاصله عدم اطمینان برای کمیت خروجی θ	یادآوری
فراوانی‌گرا	θ و μ_i همگی ثابت‌های نامعلوم هستند.	فراوانی رخداد طولانی مدت که شامل بازه θ می‌باشد.	رویکرد فراوانی‌گرای کلاسیک به یکپارچه‌سازی عدم‌قطعیت‌ها که از نظر آماری ارزشیابی نمی‌شوند، گسترش می‌یابد.
بیزی	θ و μ_i متغیرهای تصادفی هستند. توزیع احتمال آن‌ها نظرات در مورد مقادیر ورودی و خروجی کمیت‌ها را نشان می‌دهد.	بازه پوشش شامل θ بر مبنای توزیع پسین برای θ	غیرمنحصربه‌فرد بودن احتمالی بازه به‌دلیل انتخاب پیشین
اتکایی	μ_iها به‌عنوان متغیرهای تصادفی توزیع آن‌ها از مفروضات داده‌های مشاهده‌ای مورد استفاده برای برآورد μ_i و دانش کارشناسی در مورد μ_i به‌دست آمده است.	بازه پوشش شامل θ بر مبنای توزیع اتکایی برای θ	منحصربه‌فرد نبودن به‌خاطر انتخاب معادلات ساختاری

۶-۲ رویکرد بیزی

رویکرد دوم رویکرد بیزی نام دارد. این رویکرد بعد ازقضیه مبنایی که به‌وسیله توماس بیز در اواسط دهه ۱۷۰۰ [۱۲] میلادی اثبات شد، نامگذاری شده است. در این رویکرد، دانش درباره کمیت‌ها در مدل اندازه‌گیری (۱) در بند ۵ به‌عنوان مجموعه‌ای از متغیرهای تصادفی که یک توزیع احتمال مشترک برای $\mu_1,...,\mu_p$ و θ است، مدل‌سازی می‌شود. بنابراین قضیه بیز به این توزیع‌های احتمال اجازه می‌دهد تا بر اساس داده‌های مشاهده شده به‌روز شده (هم‌چنین با استفاده از توزیع احتمالات مدل‌سازی شده) و ارتباطات درونی پارامترها به‌وسیله تابع f یا مدل‌های آماری معادل تعریف شوند. بنابراین یک توزیع احتمال به‌دست می‌آید که اطلاعات θ به‌دست آمده از داده‌های مشاهده شده را توصیف می‌کند. بازه عدم‌قطعیت که شامل θ با هر احتمال مشخص است می‌تواند بنابراین از این توزیع به‌دست آید. چون اطلاعات مقادیر پارامتر از طریق توزیع احتمال توضیح داده می‌شود، روش‌های بیزی، حالت‌های احتمالی مستقیمی در مورد مقادیر θ و سایر پارامترها، با استفاده از تعریف احتمال به عنوان یک سنجش عقیده[1] ارائه می‌دهد.

1 - Belief

۶-۳ رویکرد اتکایی

۶-۳-۱ رویکرد اتکایی به‌وسیله فیشر[1] [۱۳] در سال ۱۹۳۰ میلادی توسعه یافت. در این رویکرد، یک توزیع احتمال که توزیع اتکایی نامیده می‌شود، برای θ مشروط بر داده‌های مبتنی بر روابط θ و μ_i که به‌وسیله f توصیف می‌شوند، به‌دست می‌آید و مفروضات توزیع در مورد داده‌ها برای برآورد μ_i مورد استفاده قرار می‌گیرد. بعد از به‌دست آمدن، تابع توزیع اتکایی برای θ می‌تواند برای محاسبه بازه‌های عدم‌قطعیت که شامل θ با هر احتمال مشخص است مورد استفاده قرار گیرد.

۶-۳-۲ مبحثی که برای توجیه فرآیند مورد استفاده برای به‌دست آوردن توزیع اتکایی است، با استفاده از یک مثال ساده نشان داده می‌شود. فرض کنید مقادیر به‌دست آمده با استفاده از کمیت Y، به شکل معادله $Y = \mu + Z$ توصیف شود، μ : اندازه‌ده و Z: یک کمیت مشخصی که به‌وسیله متغیر تصادفی نرمال استاندارد به‌دست می‌آید. اگر y یک مقدار مشخص از مقدار مشخص Y متناظر باشد و z یک مقدار مشخص از Z باشد، بنابراین، $\mu = y - z$. علیرغم این‌که Z قابل مشاهده نیست، اطلاعات توزیع که از آن z به‌دست می‌آید یک سری مقادیر ممکن برای μ به‌وجود می‌آورد. توزیع احتمال برای Z می‌تواند برای به‌دست آوردن توزیع احتمال برای μ مورد استفاده قرار گیرد. فرآیند تبدیل رابطه $\mu = y - z$ به رابطه $\mu = y - Z$ موردی است که مبحث اتکایی را به‌وجود می‌آورد. توزیع اتکایی برای μ، توزیع احتمال برای متغیر تصادفی $y - Z$ با ثابت y می‌باشد.

۶-۴ بحث

هنگام توضیح روش‌های متفاوت برای ارزشیابی عدم‌قطعیت تحت این رویکردهای آماری، مفروضات اساسی آن‌ها، ترکیب عدم‌قطعیت‌های به‌دست آمده با استفاده از ارزشیابی نوع A و B و تفسیر احتمالی نتایج ارزشیابی عدم‌قطعیت، مورد بحث و بررسی قرار خواهد گرفت. هم‌چنین توضیحات چگونگی ارتباط روش‌های مورد استفاده در استاندارد GUM با نتایج رویکردهای فراوانی‌گرا، بیزی و اتکایی در ادامه خواهد آمد.

۷ مثال‌ها

۷-۱ کلیات

دو مثال برای نشان دادن رویکردها ارائه داده شده است. مثال ۱ مربوط به یک کمیت فیزیکی است که مورد تصحیح در تداخل زمینه است. جدول ۲ نماد استفاده شده و زیربندهای ۷-۲ تا ۷-۴ تعریف متغیرهای مسائل ارزشیابی را نشان می‌دهد. مثال ۲، کالیبراسیون طول یک بلوک سنجه است که از پیوست H.1 استاندارد GUM گرفته شده است. از آنجایی که خیلی پیچیده است در بند ۱۱، بعد از سه روش ارزشیابی عدم‌قطعیت، بحث شده و با استفاده از مثال ۱ نشان داده می‌شود.

1 - R.A.Fisher

در بندهای بعدی سه رویکرد گفته شده، در این مثال‌ها به‌کار گرفته خواهند شد.

یادآوری- واحدهای کمیت‌های داده شده اگر در مثال‌ها مهم نباشند، داده نخواهند شد.

جدول۲- نمادهایی برای مثال ۱

کمیت	نماد
کمیت فیزیکی مورد نظر (اندازه‌ده)	θ
کمیت مشخص شده به‌وسیله روش اندازه‌گیری هنگام اندازه‌گیری قبلی (به‌عنوان مثال، مقدار مورد انتظار B) (تداخل زمینه)	β
کمیت مشخص به‌وسیله روش اندازه‌گیری هنگام اندازه‌گیری کمیت فیزیکی موردنظر (به‌عنوان مثال: مقدار مورد انتظار Y)	$\gamma = \theta + \beta$
انحراف استاندارد روش اندازه‌گیری هنگام اندازه‌گیری کمیت فیزیکی موردنظر (به‌عنوان مثال: انحراف استانداردY)	σ_Y
انحراف استاندارد روش اندازه‌گیری هنگام اندازه‌گیری زمینه (به‌عنوان مثال: انحراف استانداردB)	σ_B

۷-۲ مثال ۱- الف

پنج مقدار اندازه‌گیری شده، که به‌طور مستقل، از سیگنال به همراه زمینه به‌دست آمده‌اند، مشاهده می‌شود. در هر مقدار اندازه‌گیری شده فرض می‌شود که تحقق یک متغیر تصادفی، Y، از طریق توزیع گاوسی با میانگین $\gamma = \theta + \beta$ و انحراف استاندارد σ_Y است. مقادیر اندازه‌گیری شده،y، از سیگنال به همراه زمینه عبارت هستند از:

۳٫۸۷۴, ۳٫۶۳۷, ۲٫۹۹۴, ۳٫۴۴۲, ۳٫۷۳۸

این داده‌ها دارای میانگین نمونه‌ای $\bar{y}$ = ۳٫۵۳۷ و انحراف استاندارد نمونه‌ایs_y =۰٫۳۴۲ هستند.

به‌طور مشابه پنج مقدار اندازه‌گیری شده، به‌طور مستقل از زمینه به‌دست آمدند که فرض می‌شود این مقادیر اندازه‌گیری شده برای تحقق یک متغیر تصادفی، B، با توزیع گاوسی با میانگین β و انحراف استاندارد σ_B هستند. مقادیر b مشاهده شده، از زمینه عبارتند از:

۱٫۲۰۰, ۱٫۱۳۷, ۱٫۳۰۶, ۱٫۰۸۵, ۱٫۴۱۰

از آنجائی که مقادیر اندازه‌گیری شده برای هر کمیت، یک منبع عدم‌قطعیت است، مثال (۱- الف) یک تفسیر آماری ساده برای هر رویکرد دارد.

۷-۳ مثال ۱- ب

مثال (۱- ب) همانند مثال (۱- الف) است با این تفاوت که ارزیابی زمینه بر مبنای دانش کارشناسی یا تجربه قبلی است، نه داده‌های آزمایشی جدید. در این مورد بر این باور هستیم که زمینه β از یک توزیع

یکنواخت (یا مستطیلی) با نقاط پایانی ۱/۱۲۶ و ۱/۳۲۹ است. از آنجا که قضاوت کارشناسی اعمال می‌شود، عدم‌قطعیت مربوط به مقدار زمینه با استفاده از ارزشیابی نوع B به‌دست خواهد آمد. مثال (۱- ب) می‌تواند وضعیت اندازه‌گیری واقعی را نزدیک‌تر و بهتر از مثال (۱- الف) مورد بررسی قرار دهد.

۴-۷ مثال ۱- پ

مثال (۱- پ) شبیه مثال (۱- ب) است به استثنای این که سیگنال θ به مقدار زمینه نزدیک‌تر است. داده‌های مشاهده شده قبلی برای سیگنال به همراه زمینه در این مورد، عبارت هستند از:

۱/۱۹۲, ۱/۲۵۶, ۱/۱۱۴, ۱/۰۷۸, ۱/۳۴۰

فقط با سیگنال بالای زمینه، مثال (۱- پ) نشان می‌دهد که چگونه محدودیت‌های فیزیکی می‌توانند در ارزشیابی عدم‌قطعیت برای هر رویکرد دخیل باشند.

۸ رویکرد فراوانی‌گرا برای ارزشیابی عدم‌قطعیت

۱-۸ روش پایه

۱-۱-۸ در مفهوم فراوانی‌گرا، پارامترها ثابت‌های نامعلوم هستند. با توجه به قرارداد برای مشخص کردن متغیرهای تصادفی به‌وسیله مقادیر مشاهده شده از متغیرهای تصادفی با حروف بزرگ و مقادیر مشاهده شده متغیرهای تصادفی با حروف کوچک، بازه اطمینان می‌تواند از یک کمیت اساسی برای θ به‌دست آیند، برای مثال: یک تابع $W(Y,\theta)$ از (احتمالات چند متغیره) داده Y و پارامتر θ که توزیع احتمال آن مستقل از پارامتر است (چنین توزیعی می‌تواند تعیین شود) بنابراین، یک بازه اطمینان ٪$(1 - \alpha)$۱۰۰ برای θ می‌تواند تعیین شود که این کار به‌وسیله درصدهای پایین‌تر و بالاتر l_a و u_a برای $P_\theta(l_\alpha \le W(Y,\theta) \le u_\alpha) = 1 - \alpha$ انجام می‌گیرد.

۲-۱-۸ برای مثال، فرض کنید $Y = (Y_1, \ldots, Y_n)$ متغیرهای تصادفی هستند و به‌صورت $N(\mu,\sigma^2)$، با متغیرهای تصادفی بیشتر $\bar{Y} = \sum_{i=1}^{n} Y_i/n$ توزیع شوند. اگر پارامتر مورد نظر μ باشد بنابراین برای σ معلوم، $Z = \frac{\bar{Y} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$ یک کمیت محوری است. بازه اطمینان فراوانی‌گرا برای μ برابر است با:

$$\bar{Y} \pm \frac{\sigma}{\sqrt{n}} z_{\alpha/2}, \quad (۴)$$

که z_β بر حسب درصد β۱۰۰ از توزیع نرمال استاندارد است.

اگر σ نامعلوم نباشد، می‌تواند از طریق انحراف استاندارد نمونه برآورد شود:

$$S = \sqrt{\frac{\sum_{j=1}^{n} \left(Y_j - \bar{Y}\right)^2}{n-1}}$$

بنابراین، کمیت محوری صحیح بر μ با جایگزینی S به‌جای σ در بازه (۴) به‌صورت زیر است:

$$\frac{\bar{Y}-\mu}{S/\sqrt{n}} \sim t_{n-1}. \qquad (۵)$$

از این رو، بر اساس توزیع t- استیودنت بازه اطمینان ٪$(\alpha$ - ۱)۱۰۰ برای μ به‌صورت زیر است:

$$\bar{Y} \pm \frac{S}{\sqrt{n}} t_{n-1,1-\alpha/2},$$

که $t_{n-1,\beta}$، برابر با درصد β۱۰۰ از توزیع t با ۱- n درجه آزادی است.

۸-۱-۳ به جای کمیت‌های محوری دقیق، که فقط در موقعیت های ساده وجود دارند، کمیت‌های محوری تقریبی نیز در کاربردها به‌کار گرفته می‌شوند. برای نمونه‌های بزرگ، قضیه حد مرکزی می‌تواند برای به‌دست آوردن بازه‌های اطمینان تقریبی مبتنی بر توزیع نرمال به‌کار گرفته شود.

۸-۱-۴ روش‌های دیگری برای به‌دست آوردن بازه‌های اطمینان (معکوس کردن یک آماره آزمون، محور تابع توزیع تجمعی پیوسته، مرتب کردن مقادیر گسسته یک نمونه با توجه به احتمالات و ...) در منبع [۱۴] بحث شده‌اند. بعضی از آن‌ها در مثال ۱ ذکر شده‌اند. روش آماری مبتنی بر رایانه، که خودگردان نامیده می‌شود، هم‌چنین می‌تواند برای ایجاد بازه اطمینان برای کمیت‌های محوری که توزیع‌های نامعلوم دارند، مورد استفاده قرار گیرد. رویه خودگردان در زیربند ۸-۲ مورد بحث قرار می‌گیرد.

۸-۱-۵ اگر چه توجیهات تکراری شفاف از ملاحظات علمی بنیادی نشده است، رویه‌های توصیه شده در استاندارد GUM می‌تواند برای به‌دست آوردن بازه اطمینان تقریبی برای اندازه‌ده مورد استفاده قرار گیرد. چنین بازه‌های اطمینان بر اساس یک کمیت تقریبی محوری با یک توزیع t فرضی به‌دست آمده از مدل اندازه‌گیری (۱)، قرار دارند. با استفاده از این رویه کمیت‌های نامعلوم $\mu_1,...,\mu_p$ به‌وسیله مقادیر $x_1,...,x_p$ از اندازه‌گیری فیزیکی یا سایر منابع برآورد می‌شوند. بعضی از مقادیر x_i ممکن است میانگین‌های ساده یا دیگر توابع طراحی داده‌ها جهت برآورد کمیت‌های μ_i، $i = 1,...,m$ باشند. عدم‌قطعیت مرتبط استاندارد $u(x_i)$ نیز از داده‌ها به‌وسیله روش‌های آماری ارزشیابی می‌شود، که معمولاً انحراف استاندارد نمونه یا با استفاده از رویه‌های استوار مبتنی بر رتبه‌بندی هستند. چنین روش‌هایی به‌عنوان ارزشیابی‌های نوع A عدم‌قطعیت شناخته می‌شوند. درجه‌های آزادی v_i مرتبط با $u(x_i)$ از اندازه نمونه مورد استفاده برای برآورد μ_i تعیین می‌شود.

۸-۱-۶ از آنجائی که اندازه‌گیری فیزیکی احتمالا همیشه ممکن نمی‌باشد یا برای بعضی از مقادیر μ_iها امکان‌پذیر نیست، برآوردهای x_i از μ_i برای برخی از مقادیرi، بگویید $i = m + 1,...,p$ از طریق ارزشیابی‌های ذهنی (یا ذهنی بالقوه) به‌دست می‌آیند و همراه با x_i برای $i = 1,...,p$ مورد استفاده قرار می‌گیرند که آن‌ها نیز از ارزشیابی‌های عدم‌قطعیت نوع A به‌دست آمده‌اند. بنابراین، انواع اطلاعات غیر آماری برای برآورد $\mu_{m+1},...,\mu_p$ با استفاده از ارزشیابی عدم‌قطعیت نوع B مورد استفاده قرار می‌گیرد، که شامل قضاوت علمی، مشخصات سازندگان یا سایر اطلاعات مرتبط غیرمستقیم و یا اطلاعاتی که به‌طور ناقص مشخص شده، می‌باشد.

یادآوری- گاهی اوقات عدم‌قطعیت‌ها هم از طریق ارزشیابی نوع A و هم نوع B به‌دست می‌آیند.

۸-۱-۷ استاندارد GUM پیشنهاد می‌کند که بعضی از مدل‌های اندازه‌گیری مرتبط با اندازه‌ده θ برای وارد کردن کمیت‌های $\mu_1, \ldots, \mu_p$ برای محاسبه y از $x_1, \ldots, x_p$ مورد استفاده قرار گیرند. بنابراین، مقدار اندازه‌گیری شده (یا برآورد شده) y از θ به‌صورت:

$$y = f(x_1, \ldots, x_m, x_{m+1}, \ldots, x_p),$$

به‌دست می‌آید که Y ارزشیابی شده، $y = f(x_1, \ldots, x_p)$ مقدار اندازه‌گیری شده θ در نظر گرفته می‌شود.

۸-۱-۸ دراستاندارد GUM، قانون نشر عدم‌قطعیت برای ارزشیابی عدم‌قطعیت استاندارد، $u(y)$ مرتبط بـا y استفاده می‌شود. عدم‌قطعیت‌های استاندارد $u(x_1), \ldots, u(x_p)$ مرتبط با مقادیر $\mathrm{x} = (x_1, \ldots, x_p)$ در بسط سری تیلور[1] تابع $f(x_1, \ldots, x_p)$ در $\mu_1, \ldots, \mu_p$ استفاده می‌شود، که مرحله اول آن عبارت است از:

$$f(x_1, \ldots x_p) \approx \mathrm{f}(\mu_1, \ldots, \mu_p) + \sum_{i=1}^{p} c_i (x_i - \mu_i). \tag{۶}$$

بر مبنای $(\mu_1, \ldots, \mu_p)$ به‌وسیله μ، مشتق های جزئی:

$$c_i = \left.\frac{\partial f}{\partial \mu_i}\right|_{\mu = \mathrm{x}}$$

ضریب حساسیت نامیده می‌شود. با کاربرد قانون نشر عدم‌قطعیت در استاندارد GUM عدم‌قطعیت اسـتاندارد تقریبی مرتبط با y به شکل زیر به دست می‌آید:

$$u(y) \approx \sqrt{\sum_{i=1}^{p} c_i^2 u^2(x_i) + 2 \sum_{i<j} c_i c_j u(x_i, x_j)}, \tag{۷}$$

1 - Taylor series expansion

که در آن:

$u(x_i,x_j)$ کوواریانس بین X_i و X_j است.

۸-۱-۹ برای ارزشیابی عدم‌قطعیت استاندارد $u(y)$، در استاندارد GUM از درجه‌های آزادی مؤثر v_{eff} که از فرمول ولش ساترثویت[1] محاسبه می‌شود، استفاده می‌شود.

$$v_{eff} = \frac{u^4(y)}{\sum_{i=1}^{p} \frac{c_i^4 u^4(x_i)}{v(x_i)}} \tag{۸}$$

یادآوری- منبع [۱۵] از خاصیت غیر شهودی با توجه به مطالعات درون آزمایشگاهی استفاده می‌کند، که یک بازه اطمینان مبتنی بر تقریب ولش ساترثویت است که ممکن است برای تفاوت بین آزمایشگاهی کمتر از یکی از اجزای آن باشد.

۸-۱-۱۰ در خاتمه، به منظور ایجاد یک بازه اطمینان برای θ، کمیت محوری تقریبی برابر است با:

$$W(y,\theta) = \frac{y - \theta}{u(y)} \tag{۹}$$

به‌کار رفته است. با توجه به استاندارد GUM،

$$W(y,\theta) \sim t(v_{eff}), \tag{۱۰}$$

است که $W(Y,\theta)$ یک کمیت محوری تقریبی است که یک توزیع t با v_{eff} درجه آزادی دارد. بازه اطمینان $\%$۱۰۰(۱ - α) است.

$$y \pm u(y) t_{v_{eff}, 1-\alpha/2}, \tag{۱۱}$$

---

1 – Welch-Satterthwaite

بنابراین برای θ می‌توان ٪$(\alpha$ - ۱) ۱۰۰ را به‌عنوان بازه عدم‌قطعیت پیشنهاد کرد. نصف عرض $t_{v_{eff}, 1-\alpha/2}u(y)$ از این بازه به عنوان عدم‌قطعیت گسترش یافته مرتبط با y شناخته می‌شود.

۸-۱-۱۱ این توصیه با تمرین استانداردهای آماری سازگار است، زمانی‌که همه عدم‌قطعیت‌ها با استفاده از ارزشیابی نوع A مشخص می‌شوند، در این صورت برآورد آماری که اغلب استفاده می‌شود برای کمیت ورودی ویژه μ میانگین نمونه‌های مشاهده شده n می‌باشد. روش مرسوم برای خلاصه کردن داده‌ها جهت به‌دست آوردن عدم‌قطعیت استاندارد نوع A، برآوردگر $S/\sqrt{n}$ با درجه آزادی n-۱ است. این مسأله، بر اساس این واقعیت است که $(n\text{-}1)\ S^2/\sigma^2$ دارای یک توزیع کای‌دو با درجه آزادی n-۱ است. این روش در آماره‌های کلی از $Y = G(X_1,...,X_p)$ که در آن برآوردگرهای X_i که در آن $i = 1,...,p$ است و از قضیه حد مرکزی تبعیت می‌کند، کاربرد دارد. در واقع در این موقعیت، انحراف استاندارد Y می‌تواند به‌طور تقریبی به‌وسیله عبارت (۷) با $u(x_i,x_j)$ از طریق $\mathrm{Cov}(X_i,X_j)$ جایگزین شود.

روش استاندارد GUM مجموعه‌ای از مقیاس‌های اندازه‌شناسی را ارائه می‌دهد اما با ارائه مفروضات زیر محدود می‌شود:

- خطی بودن تابع f: به طور ایده‌آل حساسیت بهتر است زیاد نبوده و نیز نباید از بین برود؛
- نرمال بودن توزیع احتمال برآوردگرهای نقطه‌ای $Y = f(X_1,...,X_p)$: ممکن است حتی به طور تقریبی برای نمونه‌های کوچک نگه داشته شود؛
- اعتبار فرمول (۸) ولش ساترثویت: ممکن است که این فرمول زمانی‌که کمیت‌های ورودی متقابلاً مستقل هستند به خوبی کار نکند، یا اگر کمیت‌ها ورودی به طور نرمال توزیع نشده باشند و عدم‌قطعیت‌های استاندارد مشابه هم نباشند (چنانچه درجه‌های آزادی برای توزیع‌های قانون کای‌دو مرتبط نباشند، تفسیر آن‌ها مشکل خواهد شد، در واقع، آن‌ها از تئوری آماری استفاده نخواهند کرد).

۸-۱-۱۲ به‌موجب عبارت (۷) در جنبه فراوانی‌گرا، مفاهیم تئوری تصمیم آماری می‌توانند به‌کار گرفته شوند و واریانس (عدم‌قطعیت استاندارد مربع) $u^2(y)$ به‌عنوان خطای میانگین مربع برآوردگر آماری $f(x_1,x_2,...,x_p)$ تفسیر شود. این مراحل می‌توانند انجام شوند به شرط این که کمیت‌هایی که عدم‌قطعیت آن‌ها با استفاده از ارزشیابی نوع B تعیین شده‌اند، به‌عنوان مثال، $x_{m+1},....,x_p$ عمدتاً به‌وسیله انتگرال‌گیری از توزیع آن حذف شوند (به منبع ۵ مراجعه شود). اگر f به اندازه کافی به سمت خطی بودن نزدیک باشد عبارت (۷) اولین تقریب خطای میانگین مربع را ارائه می‌دهد.

۸-۱-۱۳ بحث انجام شده در مثال ۱ یک رویه فروانی‌گرای دیگر را برای به‌دست آوردن بازه اطمینان ارائه می‌دهد.

۸-۲ بازه‌های عدم‌قطعیت خودگردان

۸-۲-۱ خودگردان یک راهبرد نمونه‌گیری دوباره [۱۶] برای برآورد پارامترهای توزیع نظیر واریانس و تعیین بازه‌های اطمینان برای پارامترها زمانی است که شکل توزیع بعدی نامعلوم است. ایده کلیدی برای این روش، این است که ارتباط احتمال توزیع تجمعی(CDF) F برای Y و یک نمونه از F شبیه ارتباط بین برآورد CDF از $\widehat{F}$ است که ممکن است توزیع تجربی به‌دست آمده از نمونه نباشد و یک توزیع ثانویه باشد که از $\widehat{F}$ به‌دست آمده است. زمانی‌کهF در دسترس نباشد تابع توزیع را نمی‌تواند رسم کند، اما رایانه‌های پیشرفته این فرصت را دارند که مقادیر زیادی از نتایج را از $\widehat{F}$ به‌دست آورند. بنابراین، از نمونه اولیه برای تشکیل یک مقدار تقریبی$\widehat{F}$ از F استفاده می‌شود و بنابراین محاسبه توزیع نمونه‌ها از پارامترها بر مبنای $\widehat{F}$ برآورد می‌شوند. این محاسبه از طریق تولید بسیاری از نمونه‌های ثانویه و ایجاد برآورد (یا تابع برآورد) برای هر نمونه ثانویه انجام می‌شود. اگر $\widehat{F}$ یک مقدار تقریبی خوب برای F باشد، بنابراین H، توزیع نمونه‌ای از برآورد بر مبنای$\widehat{F}$ است که به طور کلی تقریب خوبی برای تقریب توزیع نمونه جهت برآورد بر اساسF است. H به طور معمول، توزیع خودگردان پارامتر نامیده می‌شود.

۸-۲-۲ دو رویه خودگردان مفید وجود دارد که به ترتیب برای استنباط ناپارامتریک و پارامتریک مفید هستند. خودگردان ناپارامتریک بر مبنای توزیع تجربی $\widehat{F}$ که به‌وسیله نمونه اولیه F ایجاد شده است قرار دارد. در جنبه خودگردان پارامتریک، توزیع احتمالF عضوی از خانواده پارامترهای از قبل توصیف شده است و $\widehat{F}$ از برآورد پارامترها به‌وسیله داده‌ها به‌دست می‌آید.

یادآوری- از آنجا که در مسائل اندازه‌شناسی شاخص، داده‌ها به اندازه کافی برای اطمینان از اعتبار رویکرد خودگردان ناپارامتریک، بزرگ نیستند، این رویکرد در این‌جا مورد بررسی قرار نمی‌گیرد.

۸-۲-۳ فرض کلیدی به‌کار رفته در ساختن بازه اطمینان در استاندارد GUM، فرمول (۱۰) است که ممکن است به‌طور تقریبی حتی برای مسائل جزئی و ساده نیز به‌کار گرفته نشوند. به هر حال، خودگردان، این امکان را ایجاد می‌کند که بدون در نظر گرفتن مفروضاتی مانند فرمول (۱۰) به‌دست بیایند. یک روش برای به‌دست آوردن چنین بازه‌هایی رویکرد «t- خودگردان» است. این رویه یک توزیع تجربی برای کمیت محوری تقریبی$W(Y,\theta)$ [به‌منظور جایگزینی توزیعt در فرمول (۱۰)] ایجاد می‌کند. زمانی‌که فرمول (۱۰) صحیح باشد توزیع t- خودگردان توزیع t را دوباره به‌وجود خواهد آورد. توزیع تجربی t- خودگردان، بنابراین برای ایجاد یک بازه اطمینان دقیقاً همانند روش توزیع t در ساخت فرمول (۱۱) به‌کار گرفته می‌شود.

برای ارتباط بین خودگردان کردن و روش‌های پیشنهادی در GUMS1، به زیربند ۱۲-۲ مراجعه شود.

۸-۲-۴ طرح کلی ایجاد نمونه خودگردان به شرح زیر است. فرض کنید x_1 و $u(x_1)$ میانگین و انحراف استاندارد برای متغیر تصادفی X_1 هستند که فرض می‌شود از یک توزیع احتمال متعلق به خانواده پارامتری از قبل توصیف شده، تبعیت می‌کند. در اینجا برای توضیح از توزیع گاوسی استفاده می‌شود:

الف- x_1 و $u(x_1)$ میانگین و انحراف استاندارد برآورد شده یک نمونه تصادفی با اندازه k از توزیع گاوسی است.

ب- با استفاده از $N\left(x_1,u^2(x_1)\right)$ یک نمونه با اندازه نمونه‌ای k به‌دست آورید که $\{x^*_{1,1},\ldots,x^*_{1,k}\}$ است.

پ- با استفاده از $\{x^*_{1,1},\ldots,x^*_{1,k}\}$ میانگین x^*_1 و خطای استاندارد نمونه‌ای $u(x^*_1)$ را محاسبه کنید.

$\{x^*_1,\ldots,u(x^*_1)\}$ نمونه خودگردان X_1 است. به‌طور مشابه برای عدد B داده شده، نمونه خودگردان B می‌تواند در مورد هر متغیری به‌وجود آید.

۸-۲-۵ از آن‌جا که در استاندارد GUM داریم: $\left(x_i,u(x_i)\right)$ و برای $i = 1,\ldots,p$ با ورودی y و نیز $u(y)$ و $W(Y,\theta)$ مصداق دارد ، نمونه‌های خودگردان $\{x^*_i,u(x^*_i)\}$، $i = 1,\ldots,p$ (به زیربند ۴-۲-۸ مراجعه شود) که می‌تواند به عنوان ورودی در نظر گرفته شود، y^* و $u(y^*)$ را محاسبه می‌کند:

$$W^* = W(y^*,y) = \frac{y^* - y}{u(y^*)}. \qquad (۱۲)$$

۸-۲-۶ برای به‌دست آوردن توزیع خودگردان برای $W(Y,\theta)$، که B به اندازه کافی بزرگ است، مثلاً: ۱۰۰۰۰۰، نمونه‌های خودگردان برای B به‌وجود می‌آید: $\left\{x^*_i(b),u\left(x^*_i(b)\right)\right\}$، $i = 1,\ldots,p$ و برای هر محاسبه $W^*(b)$، که $b = 1,\ldots,B$ است. درصد α ۱۰۰ اُمین توزیع t- خودگردان برای $W(Y,\theta)$ است بنابراین مقدار تقریبی $\hat{t}_\alpha$ به‌صورت زیر است:

$$\left|\{W^*(b) \le \hat{t}_\alpha\}\right|/B = \alpha,$$

که |A| تعداد مؤلفه‌ها در مورد A است. در خاتمه ٪$(۱ - \alpha)$۱۰۰ بازه اطمینان t- خودگردان این گونه است:

$$\left(y - \hat{t}_{1-\alpha/2}\cdot u(y)\ \ y + \hat{t}_{\alpha/2}\cdot u(y)\right). \qquad (۱۳)$$

درصدهای t- استیودنت در حدود صفر متقارن هستند و در نتیجه، فرمول (۱۱) باید همیشه در مورد y، متقارن و متناسب باشد. در مقابل درصدهای t- خودگردان استفاده شده درفرمول (۱۳) در مورد صفر نیز متقارن هستند که منجر به یک بازه عدم‌قطعیت متقارن در مورد y می‌شود که ممکن است توصیف دقیق‌تر از وضعیت فیزیکی در بعضی از کاربردها داشته باشد. جزئیات این فرآیند یک بازه عدم‌قطعیت ٪۹۵ را به دست می‌دهد که در الگوریتم زیر نشان داده می‌شود:

الف- برای $i = 1,\ldots,p$ با استفاده از توزیع داده‌ها برای X_i، نمونه های خودگردان B به‌دست می‌آید:

$$\left(x_i^*(1),u\left(x_i^*(1)\right)\right),\ldots,\left(x_i^*(B),u\left(x_i^*(b)\right)\right)$$

ب- برای هر نمونه، خودگردان $\left(x_i^*(b),u\left(x_i^*(b)\right)\right)$، $i = 1,\ldots,p$ و $b = 1.....B$، آن گاه $y^*(b)$، $u\left(y^*(b)\right)$ و $W^*(b) = \left(y^*(b) - y\right)/u\left(y^*(b)\right)$ را با استفاده از استاندارد GUM حساب کنید.

پ- 100α اُمین درصد توزیع t- خودگردان برای $W(Y,\theta)$ بـا مقـدار $\hat{t}_\alpha$ راماننـد $\left|\{W^*(b) \le \hat{t}_\alpha\}\right|/B = \alpha$ برآورد کنید.

ت- بازه اطمینان t- خودگردان ٪۹۵ را این گونه شکل دهید: $\left(y - \hat{t}_{0/975}\cdot u(y)\ \ y + \hat{t}_{0/025}\cdot u(y)\right)$

۸-۲-۷ نمونه‌های خودگردان می‌توانند برای جایگذاری $u(y)$ با برآورد انحراف استاندارد Y، زمانی‌که تقریب تیلور فرمول (۶) نامناسب انگاشته شود، مورد استفاده قرار گیرد. بـرای انجـام ایـن کـار، بـرای $i= 1,\ldots,p$ و $b=1,\ldots,B$، تنهـا ورودی بـرآوردی $x_i^*(b)$ ایجـاد شـده اسـت. بـرای هـر نمونـه خودگردان، $y^*(b) = f\left(x_1^*(b),\ldots,x_p^*(b)\right)$ ارزشیابی شده است. با برآورد خودگردان عـدم‌قطعیـت اسـتاندارد مرتبط با y، انحراف استاندارد نمونه B به‌صورت زیر به‌دست می‌آید:

$$u_c(y) = \sqrt{\sum_{b=1}^{B}[y^*(b) - y^*(.)]^2\Big/(B-1)}, \qquad y^*(.) = \sum_{b=1}^{B} y^*(b)\Big/B.$$

۸-۲-۸ در خاتمه، زمانی‌که تقریب تیلور ممکن است مناسب نباشد و به‌طور معنی‌داری تقارن چشـمگیری در توزیع Y به چشم آید، یک خودگردان آشیانی، $B_1 \times B_2$ نمونه‌های خودگردان را می‌توان برای ساخت یک بازه خودگردان با استفاده از برآوردگر انحراف استاندارد t– خودگردان انجام داد. B_1 نمونه‌هـای خـودگردان برآورد داده‌های ورودی است و y^* متناظر تولید می‌شود. برای هر نمونه خودگردان، $u_c(y^*)$ ازطریق نمونه‌های خودگردان سطح دوم B_2 محاسبه می‌شود، و

$$\frac{y^* - y}{u(y^*)}$$

ارزشیابی می‌شود. مجموعه B_1 هم‌چنین نسبت‌هایی برای برآورد درصد $W(Y,\theta)$ مورد استفاده قرار می‌دهد، که منجر به ایجاد بازه‌های t– خودگردان همان‌طورکه در فرمول (۱۳) دیدیم می‌شـود. یـک الگـوریتم بـرای ایجاد بازه عدم‌قطعیت ٪۹۵ با استفاده از خودگردان آشیانی به‌صورت زیر است:

الف- برای $i = 1,\ldots,p$، با استفاده از توزیع برای X_i، B_1 به‌وجود می‌آید و نمونه‌های خودگردان سـطح اول به شکل $x_i^*(1),\ldots,x_i^*(B_1)$ به‌دست می‌آید.

ب- برای هر نمونه خودگردان سطح اول $x_i^*(b)$، $i = 1,\dots,p$، $b_1= 1,\dots,B_1$، آن‌گاه محاسبه کنید: $y^*(b_1)=f\left(x_1^*(b_1),\dots,x_p^*(b_1)\right)$، و $W^*(b_1) = \left(y^*(b_1)-y\right)/u\left(y^*(b_1)\right)$ که $u\left(y^*(b_1)\right)$ از طریق خودگردان سطح دوم و الگوریتم زیر تعیین می‌شود:

۱- برای $i=1,\dots,p$، با استفاده از توزیع برای μ_i خودگردان سطح دوم، B_2 و نمونه‌های $x_i^*(1),\dots,x_i^*(B_2)$ به‌دست می‌آید.

۲- برای نمونه خودگردان سطح دوم، $y^*(b_2) =f\left(x_1^*(b_2),\dots,x_p^*(b_2)\right)$ را ارزشیابی کنید.

۳- برآورد خودگردان از عدم‌قطعیت استاندارد $y^*(\mathrm{b}_1)$ را همانند انحراف استاندارد نمونه این گونه شکل دهید :

$$\mathrm{u}\left(y^*(b_1)\right) = \sqrt{\sum_{b_{2=1}}^{B_2}[y^*(b_2) - y^*(.)]^2 \Big/ (B_2 - 1)}$$

که در آن از تکرارهای B_2، $y^*(.) = \sum_{b_{2=1}}^{B_2} y^*(b_2)/B_2$ به‌دست آمده است.

پ- برآورد ۱۰۰α اُمین درصد توزیع t– خودگردان از $W(Y,\theta)$ با مقدار $\hat{t}_\alpha$ برابر $\left|\{W^*(b_1) \leq \hat{t}_\alpha\}\right|/B_1= \alpha$ به‌دست می‌آید.

ت- بازه اطمینان«٪۹۵ t- خودگردان آشیانی این گونه است»:

$$\left(y - \hat{t}_{0,975}\cdot u(y),\ y + \hat{t}_{0,025}\cdot u(y)\right).$$

اگر چه این رویکرد کلی‌تر می‌باشد، خودگردان آشیانی بیشتر مورد محاسبه قرار می‌گیرد و اجرای آن سخت‌تر است. با این وصف روش خودگردان ساده‌تری در تحلیل تمام مثال‌ها انتخاب شد.

۸-۳ مثال ۱

۸-۳-۱ کلیات

۸-۳-۱-۱ به‌عنوان یک توضیح، مدل آماری داده شده برای مثال ۱ را در بند ۷ در نظر بگیرید،

$$Y_i = \theta + \beta + \varepsilon_i, \qquad i = 1,\dots,n, \qquad (۱۴)$$

که θ یک‌اندازه‌ده است، β نشانگر زمینه و ε_i خطاهای مستقل $N(0,\sigma^2)$ هستند. برای یک مقدار ثابت β، با γ نشانگر میانگین داده‌ها، معادله سنجش برای این مدل $\theta = f(\beta,\gamma)= \gamma - \beta$ است.

۸-۳-۱-۲ اگر زمینه، β، توزیع یکنواخت در بازه $(\mathrm{a{-}d} \quad \mathrm{a{+}d})$ باشد، بازه برای θ مشتق شده با استفاده از استاندارد GUM برابر است با:

$$\bar{Y} - a \pm 2\sqrt{\frac{\sigma^2}{n} + \frac{d^2}{3}}\ .$$

منبع[۵] خصوصیات چنین بازه‌هایی را بحث می‌کند و آن‌ها را با بازه مقایسه می‌کند.

$$\bar{Y} - a\ \pm \left[2\sqrt{\frac{\sigma^2}{n}} + d\right], \tag{۱۵}$$

که به‌وسیله آیزن‌هارت[1] پیشنهاد شده است [۱۷] و می‌تواند به صورت زیر برانگیخته شود. از آنجا که توزیع شرطی $\bar{Y}$ برای β ، گاوسی است، $N(\theta + \beta, \sigma^2 / n)$ ،

$$\mathrm{p}\left(|\bar{Y} - \theta - \beta| \le \frac{2\sigma}{\sqrt{n}}\right) \ge 0{,}95,$$

درحالی‌که

$$\mathrm{P}(|\alpha - \beta| \le\ d) = 1.$$

با استفاده از بازه آیزن‌هارت در فرمول(۱۵) آن‌گاه،

$$\mathrm{P}\left(|\bar{Y} - \alpha - \theta| \le \frac{2\sigma}{\sqrt{n}} + d\right) \ge 0{,}95, \tag{۱۶}$$

به‌دست می‌آید.

۸-۳-۱-۳ به هر حال، اگر $d > 12\sigma/\sqrt{n}$ ، بازه پیشنهاد شده در استاندارد GUM کـه حـاوی بـازه (۱۵) است، تفاوت بین دو رویکرد را نشان می‌دهد.

۸-۳-۱-۴ بازه (۱۵) می‌تواند برای نسبت $\sqrt{n}\left(\bar{Y} - a - \beta\right)/S$ که از توزیع t، تبعیت می‌کند، تنظـیم شـود. هم‌چنین می‌تواند برای سایر توزیع‌ها برای زمینه مورد نظر (مثلثـی، ذوزنقـه‌ای و غیـره) نیـز تنظـیم شـود. روش‌های فراوانی‌گرای متفاوت برای ایجاد بازه اطمینان در این موقعیت وجود دارند. در واقع، در مدل (۱۴) $\bar{Y}$ زیرمجموعه تمام اطلاعات را در مورد داده θ استنتاج می‌کند (که $\bar{Y}$ یک آمـاره بسـنده بـرای θ اسـت) کـه چگالی احتمال آن از قرار زیر است:

$$\frac{\sqrt{n}}{2\sqrt{2\pi}\sigma d}\int_{a-d}^{a+d} e^{-0.5n(\bar{Y} - \theta - \beta)^2/\sigma^2}\,\mathrm{d}\beta.$$

1 -Eisenhart

شکل خاص این توزیع منجر به تولید به بازه‌های اطمینان مختلف می‌شود تا محاسبه شوند (به‌طوری که برآوردگر احتمالی ماکسیمم درست‌نمایی حداکثر $\bar{Y} - a$، با طول‌های متفاوت متمرکز شده‌اند)[۱۴].

۸-۳-۲ مثال ۱- الف

مثال ساده داده شده دربند ۷ مقادیر اندازه‌گیری شده در مدل (۱۴) را با $\bar{y}$ = ۳٫۵۳۷ و $u(\bar{y})$ = ۰٫۱۵۳ خلاصه می‌کند. بعد برای $\sigma/\sqrt{n}$ در نامعادله (۱۶) و عامل۲ بهتر است با درصد t- توزیع و با درجه آزادی مؤثر ۵٫۱۵ جایگزین شود. در مثال (۱- الف)، زمینه β می‌تواند از مقادیر اندازه‌گیری شده همانطوری که از توزیع گاوسی نتیجه شده است، برآورد شود که منجر به $\bar{b}$ = ۱٫۲۲۸ و $u(\bar{b})$ = ۰٫۰۵۹ می‌شود. برآورد منتج از θ، $\bar{y} - \bar{b}$ = ۲٫۳۰۹ با عدم‌قطعیت استاندارد $\sqrt{u^2(\bar{y}) + u^2(\bar{b})}$ = ۰٫۱۶۴ است. بازه اطمینان استاندارد GUM نیز

$$۲٫۳۰۹ \pm ۲٫۵۴۸ \times ۰٫۱۶۴ = ۲٫۳۰۹ \pm ۰٫۴۱۷ = (۱٫۸۹۲ \quad ۲٫۷۲۷)$$

٪$(۱ - \alpha)$۱۰۰ بازه اطمینان t– خودگردان با توجه به فرمول (۱۳) برابر با $\left(2{,}309 + 0{,}164 \cdot \hat{t}_{\alpha/2} \quad 2{,}309 - 0{,}164 \cdot \hat{t}_{1-\alpha/2}\right)$ می‌باشد که $\hat{t}_\beta$ برابر β۱۰۰ اُمین درصد W^* در فرمول (۱۲) است.

برای بهره‌مندی استفاده‌کنندگان از زبان R و WinBUGS، بعضی از اجزای کد R [۱۸] و WinBUGS [۱۹] برای ارائه مثال‌هایی در مورد برخی از مفاهیم در این استاندارد مورد بررسی قرار می‌گیرد. برای مثال (۱- الف) و یک برنامه R برای ایجاد B = 10000 برای تفهیم W^* به شکل زیر فهرست شده است.

```
B = 10000
y.star = rnorm(B, mean=3.537, sd=0.153)
u.y.star = 0.153 * sqrt(rchisq(B, df=4)/4)
b.star = rnorm(B, mean=1.228, sd=0.059)
u.b.star = 0.059 * sqrt(rchisq(B, df=4)/4)
w.star = ((y.star-b.star)-2.309)/sqrt(u.y.star^2+u.b.star^2)
```

بازه اطمینان ٪۹۵ t– خودگردان برمبنای چندک‌های[۱] ۰٫۰۲۵ و ۰٫۹۷۵ از توزیع شبیه سازی‌شده مطابق با زیر است:

```
2.309 - quantile(w.star, c(0.975,0.025))*0.164
## 1.895754 2.7288172)
```
[۲]

1 - Quantiles

۲ - در نمونه‌های مورد نظر که با استفاده از WinBugs ، R یا سایر بسته‌های نرم‌افزارها محاسبه شده‌اند، مقادیر خروجی با استفاده از فرمت استاندارد نرم‌افزار گزارش می‌شود. همان‌طوری که با استفاده از مقادیر عدم‌قطعیت گزارش شده، واضح است، تمام ارقام خروجی می‌توانند اعداد معنی‌دار نباشند. هم‌چنین توجه کنید که خروجی استاندارد از این بسته‌های نرم‌افزاری از دوره به جای ویرگول به‌عنوان یک نشانگر اعشاری استفاده می‌کند.

برای مثال بازه اطمینان ٪۹۵ t– خودگردان با (۱٫۸۹۶ ۲٫۷۲۹)[1] داده شده است.

۸-۳-۳ مثال ۱- ب

زمانی‌که هیچ داده آماری در مورد زمینه وجود ندارد، فرض می‌شود که β یک توزیع یکنواخت در بازه (۱٫۱۲۶ ۱٫۳۲۹) دارد. بنابراین، بازه اطمینان تقریبی که با استفاده از استاندارد GUM به‌دست آمده است برابر است با:

$$3{,}537 - 1{,}228 \pm 2{,}533\sqrt{\frac{0{,}342^2}{5}+\frac{0{,}102^2}{3}} = 2{,}310 \pm 0{,}415 = (1{,}895 \quad 2{,}724).$$

بازه اطمینان آیزن‌هارت گسترده‌تر است، به‌عنوان مثال:

$$3{,}537-1{,}228\pm\left[2{,}776\frac{0{,}342}{\sqrt{5}}+0{,}102\right]=2{,}310\pm0{,}526=(1{,}783 \quad 2{,}836)$$

مانند مثال (۱- الف)، یک بازه اطمینان t– خودگردان می‌تواند برای θ ایجاد شود. برای این مثال، برآوردها و عدم‌قطعیت‌های استاندارد مرتبط برای y، β و θ از لحاظ عددی مشابه مثال (۱- الف) است. با این تفاوت که β بر اساس تجربه یا نظریات کارشناسی است و عدم‌قطعیت مرتبط با آن به‌وسیله ارزشیابی نوع B به‌دست می‌آید. بنابراین تحقق و تفهیم W^* با یک روش متفاوت از مثال (۱- الف) به‌دست می‌آید. به‌عنوان مثال، ایجاد نمونه خودگردان b^* و عدم‌قطعیت مربوط به آن. نمونه خودگردان b^* اکنون از بازه یکنواخت و معلوم (۱٫۱۲۶ ۱٫۳۲۹) با عدم‌قطعیت استاندارد ۰٫۰۵۹ به‌دست می‌آید. کُد R برای به‌دست آوردن B = ۱۰۰۰۰ و تحقق W^* به‌صورت زیر است.

```
B = 10000
y.star = rnorm(B, mean=3.537, sd=0.153)
u.y.star = 0.153 * sqrt(rchisq(B, df=4)/4)
b.star = runif(B, min=1.126, max=1.329)
u.b.star = 0.059
w.star = ((y.star-b.star)-2.309)/sqrt(u.y.star^2+u.b.star^2)
```

بازه اطمینان ٪۹۵ t- خودگردان بر مبنای چندک‌های ۰٫۰۲۵ و ۰٫۹۷۵ از توزیع به شکل زیر است.

```
2.309 – quantile (w.star, c(0.975,0.025))*0.164
## 1.918643 2.699749
```

یعنی، بازه اطمینان ٪۹۵ t- خودگردان به‌وسیله (۱٫۹۱۹ ۲٫۷۰۰) داده می‌شود.

۱- مقادیر تا سه رقم معنی‌دار در بسط عدم‌قطعیت بسط یافته، گرد شده‌اند. برای روش‌های مونت کارلو، محاسبه دوباره مثال‌ها مربوط به خطای تصادفی ناشی از شبیه‌سازی مربوط خواهد بود.

۸-۳-۴ مثال ۱- پ

از آنجایی که $\bar{y}$ = ۱٫۱۹۶، $s_{\bar{y}}$ = ۰٫۰۴۷ هر دو فاصله، نقطه پایانی منفی کمتری دارند. اگر میانگین θ معلوم مثبت باشد این نقطه پایانی با صفر جایگزین می‌شود که منجر به بازه توصیه شده در استاندارد GUM (۰٫۱۲۴ ۰) می‌شود و بازه (۰٫۲۰۲ ۰) آیزن‌هارت به‌دست می‌آید.

یک برنامه R برای به‌دست آوردن B = ۱۰۰۰۰ برای تحقق W^* جهت به‌دست آوردن بازه خودگردان مانند مثال (۱- ب) با $\bar{y}$ = ۱٫۱۹۶ و $u(\bar{y})$ = ۰٫۰۴۷ می‌باشد.

```
B = 10000
y.star = rnorm(B, mean=1.196, sd=0.047)
u.y.star = 0.047 * sqrt(rchisq(B, df=4)/4)
b.star = runif(B, min=1.126, max=1.329)

u.b.star = 0.059
w.star=((y.star-b.star)+0.032)/sqrt(u.y.star^2+u.b.star^2)
```

بازه اطمینان ٪۹۵ خودگردان به‌صورت کامل به‌صورت زیر است:

```
0.032-quantile(w.star,c(0.975,0.025))*0.075
## -0.1762648 0.1128422
```

برای مثال، بازه اطمینان خودگردان٪۹۵ با (۰٫۱۱۳ ۰٫۱۷۶-) داده می‌شود.

از آنجایی که θ معلوم مثبت است بازه اطمینان t– خودگردان ٪۹۵ به‌طور ناقص برای θ (۰٫۱۱۳ ۰) می‌باشد.

۹ رویکرد بیزی برای ارزشیابی عدم‌قطعیت

۹-۱ روش پایه

۹-۱-۱ در اندازه‌شناسی، اندازه‌ده‌ها و متغیرهای ورودی مدل (۱)،کمیت‌های فیزیکی با مقادیر کیفی ثابت هستند. با این وجود، با توجه به رویکرد بیزی، پارامترهای متناظر μ_i و θ به‌عنوان متغیرهای تصادفی در نظر گرفته می‌شوند که توزیع‌های احتمال آن‌ها اطلاعات مربوط به این کمیت‌ها را خلاصه می‌کند.

۹-۱-۲ چارچوب بیزی از تعریف احتمال که به توزیع احتمال اجازه می‌دهد بدون داده‌های فیزیکی تعریف شوند، استفاده می‌کند. برای مثال، استفاده از ویژگی‌های سازندگان یا دانش کارشناسی. در کاربردهای اندازه‌شناسی شاخص، با این‌حال مقادیر(داده‌های) اندازه‌گیری شده از کمیت‌های فیزیکی وجود دارند که می‌توانند برای برآورد یک یا چند کمیت ورودی مورد استفاده قرار گیرند. در چنین مواردی، تابع چگالی احتمال می‌تواند برای کمیت با استفاده از قضیه بیزی به‌صورت زیر به‌دست آید. فرض کنید $P(\mu_i)$ یک تابع چگالی احتمال برای μ_i باشد، مانند قبل نتایج داده‌های فیزیکی به‌دست می‌آیند. این تابع، چگالی پیشین برای μ_i نامیده می‌شود. اجازه دهید Y یک متغیر تصادفی باشد که برای تحقق y (داده) وجود دارد. چگالی احتمال $\mathrm{P}(y \mid \mu_i)$ از Y یک مدل آماری نامیده می‌شود. تحت چارچوب بیزی، از آنجائی‌که μ_i یک متغیر تصادفی است، نماد این واقعیت را بیان می‌کند که چگالی احتمال Y روی μ_i شرطی (یا وابسته) است. برای تحقق $\mathrm{P}(y \mid \mu_i)$ از Y، تابع μ_i به‌دست آمده، تابع درست‌نمایی نامیده می‌شود. با به‌کارگیری قضیه بیزی داریم:

$$\mathrm{p}(\mu_i \mid y) = \frac{p(y \mid \mu_i)p(\mu_i)}{\int p(y \mid \mu_i)\mathrm{p}(\mu_i)d\mu_i} \qquad (۱۷)$$

که چگالی پسین μ_i است که دانش ما را در مورد μ_i بعد از داده‌های مشاهده شده y_i خلاصه می‌کند.

۹-۱-۳ زمانی که از μ_i هیچ‌گونه دانش و اطلاع پیشین وجود نداشته باشد بنابراین از یک توزیع پیشین که به اصطلاح غیراطلاعاتی نامیده می‌شود [۲۰]، استفاده می‌کنیم. در مواردی که اطلاع و دانش پیشین وجود دارد از طریق یک توزیع احتمال اطلاعاتی ارائه و نشان داده می‌شود. این یکی از مکانیزم‌های تحت رویکرد بیزی است که شامل اطلاعاتی است که برای اجرای ارزشیابی عدم‌قطعیت نوع B مورد استفاده قرار می‌گیرد. شکل تابع درست‌نمایی معمولاً بر مبنای دانش فرآیند که داده‌ها را به‌وجود می‌آورد انتخاب می‌شود.

۹-۱-۴ شکل تابع درست‌نمایی و چگالی‌های پیشین، شکل چگالی پسین را تعیین می‌کند. انتخاب تابع درست‌نمایی و چگالی‌های پیشین مهم است و تا بتوان تحلیل حساسیت نتایج را با توجه به تغییرات قابل قبول در این توابع انجام داد. برای توزیع‌های پیشین، آن به این معنی است که باید نتایج استفاده از چندین چگالی مختلف را با هم مقایسه کرد. یک آزمایش متناسب تابع درست‌نمایی (مدل آماری که داده‌های اندازه‌گیری را توصیف می‌کند) شکلی از اعتبار مدل است [۲۱] که به‌طور مساوی، مدل‌های بیزی، فراوانی‌گرا و اتکایی را به‌کار می‌گیرد.

۹-۱-۵ تعریف عدم‌قطعیت ارائه داده شده در مقدمه می‌تواند در متن آماره‌های بیزی، همان‌طوری که در توزیع احتمال پسین که برای اندازه‌ده θ گفته شد، تفسیر شود. عدم‌قطعیت استاندارد، انحراف استاندارد متغیر تصادفی (کمیت) مشخص شده با توزیع احتمال است. برای به‌دست آوردن این انحراف استاندارد لازم است که ابتدا توزیع احتمال مشترک μ_i را پیدا کرد و بنابراین فرمول تغییر متغیرها را جهت استخراج توزیع برای θ به‌کار برد [۱۴]. گام‌های این توزیع می‌تواند به‌صورت ساده‌تر به‌شکل زیر باشد. برای یک تابع $h(\theta)$ مقدار مورد انتظار واریانس $\mathrm{E}\big(h(\theta)\big) = \int \dots \int h\Big(f\big(\mu_1,\dots,\mu_p\big)\Big)\, p\,\big(\mu_1,\dots,\mu_p\big)\, d\mu_1 \dots d\mu_p$ به‌دست می‌آید. واریانس متناظر می‌تواند به‌صورت $\mathrm{Var}(\theta) = E\big(\theta^2\big) - \big[\,E(\theta)\big]^2$ به‌دست آید. اغلب پیوستگی لازم با استفاده از روش‌های مونت کارلو انجام می‌شود [۲۰].

۹-۱-۶ زمانی‌که μ_iها متغیر تصادفی مستقل هستند، توزیع احتمال مشترک آن‌ها محصول توزیع‌های انفرادی است. در بسیاری از موقعیت‌ها μ_iها همانند زمانی‌که توزیع احتمال برای Y یک تابع از μ_1 و μ_2 باشد، مستقل نیستند که $\mathrm{p}\big(y \mid \mu_1,\mu_2\big)$ مدل آماری و $p\,(\mu_1,\mu_2) \neq \mathrm{p}\,(\mu_1)\,\mathrm{p}\,(\mu_2)$ هستند. بنابراین چگالی پسین برای $\big(\mu_1,\mu_2\big)$ به‌شکل زیر به‌دست می‌آید.

$$\mathrm{p}\big(\mu_1,\mu_2 \big|\, y\big) = \frac{p\big(y \mid \mu_1,\mu_2\big)p\big(\mu_1,\mu_2\big)}{\int p\big(y \mid \mu_1,\mu_2\big)p\big(\mu_1,\mu_2\big)d\mu_1 d\mu_2}$$

۹-۱-۷ یک موقعیت رایج که منجر به چنین وابستگی می‌شود زمانی است که مدل آماری یک تابع از θ و همین‌طور از بعضی از μ_iها باشد. هر دو مثال بررسی شده در این‌جا، در این طبقه‌بندی قرار می‌گیرند که نشانگر این است که تحت رویکرد بیزی هر زمانی که داده‌های اندازه‌گیری شده در دسترس باشند فرآیند تعیین توزیع‌های احتمال مربوط نیازمند یک تعریف مناسب از یک مدل آماری است. انجام این کار به‌طور خودکار منجر به توابع درست‌نمایی موردنیاز برای کاربرد قضیه بیزی و چگالی‌های پسین مناسب خواهد شد. فرآیند می‌تواند به‌طور خلاصه به‌صورت زیر بیان شود.

الف- تمام داده های اندازه‌گیری مربوط به کمیت‌های فیزیکی را معرفی کنید (پارامترها).

ب- یک مدل آماری که داده‌ها (مدل مشاهده نیز نامیده می‌شود) را به پارمترها مربوط می‌کند مشخص کنید که می‌تواند μ_iها یا بعضی وقت‌ها اندازه‌ده θ باشد.

پ- توزیع پیشین را برای تمام پارامترهای موجود مشخص کنید.

ت- قضیه بیزی را برای به‌دست آوردن توزیع‌های پسین برای پارامترها به‌کار ببرید.

ث- میانگین پسین و انحراف استاندارد پسین را برای اندازه‌ده محاسبه کنید.

ج- تحلیل حساسیت نتایج را با توجه به تغییرات ممکن در توزیع‌های پیشین اجرا کنید.

۹-۱-۸ اگر مناسب باشد، احتمالاً تقریب سری‌های تیلور و یک فرضیه نرمال جهت اجتناب از محاسبات عددی مورد استفاده قرار خواهد گرفت. به‌طور خاص، بسط سری‌های تیلور از $f(\mu_1,...,\mu_p)$ در مورد مقادیر مورد انتظار μ_i همراه با فرضیه نرمال می‌تواند برای بیان این‌که $f(\mu_1,...,\mu_p)$ تقریبی به صورت $N(f(E(\mu_1),...,E(\mu_p)),\omega^2)$, توزیع می‌شود، مورد استفاده قرار گیرد که

$$\omega^2 = \sum_{i=1}^{p} c_i^2 \,\mathrm{Var}(\mu_i) + 2\sum_{i<j} c_i c_j \,\mathrm{Cov}(\mu_i,\mu_j).$$

(μ_i,μ_j) به معنی کوواریانس از μ_i و μ_j است و c_i مشتق جزئی θ با در نظر گرفتن این‌که μ_i در مقادیر موردانتظار از μ_i ارزشیابی می‌شود، هستند.

یادآوری- فرمول‌های مشابه (۶) و (۷) در زیربند ۸-۱-۸ از بسط برای پیدا کردن برآوردی از واریانس و آنچه که θ را برآورد می‌کند استفاده می‌شود نه خود θ .

۹-۲ مثال ۱

۹-۲-۱ کلیات

اکنون این فرآیند در مثال ۱ بند ۷ نشان داده می‌شود. اندازه‌ده در این مثال به‌وسیله θ مشخص می‌شود. مدل اندازه‌گیری همان‌طوری که در زیربند ۸-۳-۱-۱ توضیح داده شد این گونه است:

$$\theta = \gamma - \beta. \qquad (۱۸)$$

۹-۲-۲ مثال ۱- الف

۹-۲-۲-۱ دو مجموعه داده مرتبط وجود دارد: (i) پنج مقدار اندازه‌گیری شده y_i که به‌طور مستقل به‌دست آمده‌اند، با سیگنال توأم با زمینه و (ii) پنج مقدار اندازه‌گیری شده b_i که به‌صورت مستقل با استفاده از زمینه به‌دست آمده‌اند. هر مقدار در مجموعه داده‌ها (i) به‌عنوان تحقق یک مفهوم از متغیر تصادفی Y_i که دارای توزیع گاوسی با میانگین $\gamma = \theta + \beta$ و انحراف استاندارد σ_Y است، در نظر گرفته می‌شود و متشابهاً برای هر مقدار در (ii) اما برای متغیر تصادفی B_i با میانگین β و انحراف استاندارد σ_B برقرار است. بنابراین مدل آماری برای Y_i این چنین است:

$$Y_i \mid \theta, \beta, \sigma_Y^2 \sim N\left(\theta + \beta, \sigma_Y^2\right),$$

و از آنجائی‌که پنج مقدار اندازه‌گیری شده به‌طور مستقل به‌دست آمده‌اند،

$$\mathrm{p}\left(y_1, \ldots, y_5 \mid \theta, \beta, \sigma_Y\right) = \left(\frac{1}{\sigma_Y \sqrt{2\pi}}\right)^5 \exp\left\{\frac{\sum_{i=1}^{5}\left(y_i - \theta - \beta\right)^2}{2\sigma_Y^2}\right\}$$

۹-۲-۲-۲ مدل آماری برای B_i عبارت است از:

$$B_i \mid \beta, \sigma_B^2 \sim N\left(\beta, \sigma_B^2\right)$$

که:

$$\mathrm{p}(b_1, \ldots, b_5 \mid \beta, \sigma_B) = \left(\frac{1}{\sigma_B \sqrt{2\pi}}\right)^5 \exp\left\{-\frac{\sum_{i=1}^{5}(b_i - \beta)^2}{2\sigma_B^2}\right\}$$

۹-۲-۲-۳ ازآنجایی که دو مجموعه از مشاهدات به‌طور متقابل مستقل هستند، مدل آماری برایY وB بدین صورت است:

$$\mathrm{p}(y, b \mid \theta, \beta, \sigma_Y, \sigma_B) = \mathrm{p}(b_1, \ldots, b_5 \mid \beta, \sigma_B)p\left(y_1, \ldots, y_5 \mid \theta, \beta, \sigma_y\right)$$

۹-۲-۲-۴ θ، β، σ_Y و σ_B چهار پارامتری هستند که به توزیع‌های پیشین اختصاص داده می‌شوند. در این مثال، هیچ‌گونه اطلاعات اضافی در مورد این پارامترها به‌جز این‌که آن‌ها غیر منفی هستند، وجود ندارد و بنابراین متغیرهای تصادفی، مستقل در نظر گرفته می‌شوند. مطلوب است که توابع پیشین اثرحداقلی بر روی نتایج تحلیل داشته باشند. چنین اثری با استفاده از اصطلاح اولویت‌های منبع به‌دست می‌آیند [۲۰]. برای پارامترهای مرتبط با میانگین‌ها، که θ و β هستند، چنین چگالی می‌تواند به‌طور تقریبی به‌صورت زیر باشد:

$$\theta \sim \text{Uniform}(0, c), \qquad \beta \sim \text{Uniform}(0, c)$$

که مقدار بزرگ برای c در نظر گرفته می‌شود. برای پارامترهای مقیاس σ_Y و σ_B، چگالی‌های پیشین منبع:

$$p(\sigma_Y) - 1/\sigma_Y\,, \qquad P(\sigma_B) = 1/\sigma_B$$

ناسره[1] هستند که به‌صورت یکپارچه و واحد نیستند. از آنجائی‌که این جنبه می‌تواند در محاسبات عددی باعث مشکلاتی شود یک چگالی مناسب مانند:

$$\sigma_Y \sim \text{Uniform}(0, \text{c}),$$

یا:

$$\sigma_Y \sim \text{Gamma}(\text{c}, \text{c})$$

با مقادیر بزرگ برای c استفاده می‌شود. نماد $Gamma(\emptyset_1, \emptyset_2)$ بیانگر توزیع گاما با پارامترهای $\emptyset_1$ و $\emptyset_2$ است که برای متغیر تصادفی X چگالی احتمال به صورت

$$\text{p}(\text{ x} \mid \emptyset_1, \emptyset_2) = \frac{\emptyset_2^{\emptyset_1}}{\Gamma(\emptyset_1)} x^{\emptyset_1 - 1} e^{-x\emptyset_2}.$$

داده می‌شود. که مشخصات و ویژگی‌های توزیع‌های پیشین را تکمیل می‌کند.

1 - Improper

۹-۲-۲-۵ کاربرد قضیه بیزی و نتایج آن در چگالی پسین مشترک برای θ ، β ، σ_Y و σ_B به‌صورت زیر است:

$$\mathrm{p}(\theta,\beta,\sigma_Y,\sigma_B|\,y,b)=\frac{p(y,b\mid\theta,\beta,\sigma_Y,\sigma_B)p(\theta)p(\beta)p(\sigma_Y)p(\sigma_B)}{\int p(y,b\mid\theta,\beta,\sigma_Y,\sigma_B)p(\theta)p(\beta)p(\sigma_Y)p(\sigma_B)d\theta\;d\beta\;d\sigma_Y\;d\sigma_B}.$$

چگلی پسین برای اندازه‌ده θ به وسیله انتگرال زیر به‌دست می‌آید:

$$\mathrm{p}(\theta\mid y,b)=\int p(\theta,\beta,\sigma_Y,\sigma_B\mid y,b)d\beta\;d\sigma_Y\;d\sigma_B.$$

توزیع پسین تمام اطلاعات در مورد θ را قبل از این‌که مقادیر اندازه‌گیری شده به‌دست بیایند خلاصه می‌کند. این توزیع به‌عنوان یک برآورد از کمیت فیزیکی و انحراف استاندارد آن به‌عنوان عدم‌قطعیت استاندارد مرتبط با این برآورد در نظر گرفته می‌شود. به‌دست آوردن بازه پوششی برای اندازه‌ده از این توزیع آسان است. این بازه پوششی یک بازه از مقادیر ممکن برای θ با یک احتمال ثابت است. در آماره‌های بیزی این بازه یک بازه معتبر نامیده می‌شود. در بسیاری از موارد هنگام کاربرد قضیه بیزی از روش‌های عددی برای انجام یکپارچگی لازم استفاده می‌شود. یک روش ممکن برای نتیجه‌گیری از توزیع‌های پسین زنجیره مارکوف و مونت کارلو [۲۲] (MCMC)[1] با استفاده از نرم‌افزارWinBUG است [۱۹]. کد این مثال، با توزیع‌های پیشین یکنواخت و C= ۱۰۰ به‌شکل زیر است:

```
Example1a{
theta~dunif(0,100)
beta~dunif(0,100)
gamma <- theta+beta
sigma.Y~dunif(0,1)
sigma.B~dunif(0,1)
tau.Y <- 1/(sigma.Y*sigma.Y)
tau.B <- 1/(sigma.B*sigma.B)
for(i in 1:n){
y[i]~dnorm(gamma,tau.Y)
b[i]~dnorm(beta,tau.B)}
}
```

با داده‌های داده شده در زیربند ۷-۲ برای $n = 5$، برنامه یک میانگین پسین از θ برابر ۲٫۳۰۹ و یک انحراف استاندارد پسین برابر با ۰٫۲۴۷ را تولید کرد. بازه اطمینان ٪۹۵ برای θ (۲٫۸۱۵ ۱٫۸۰۵) است. تحلیل حساسیت بیزی با در نظر گرفتن تغییرات در شکل چهار توزیع پیشین می‌تواند با مقادیر متفاوت c (به زیربند ۹-۲-۲-۴ مراجعه شود) و با جایگذاری خطوط

1 -Markov Chain Monte Carlo

```
tau.Y~dgamma(1,0E-5,1,0E-5)
tau.B~dgamma(1,0E-5,1,0E-5)
```

برای چهار خط

```
sigma.Y~dunif(0,1)
sigma.B~dunif(0,1)
tau.Y <- 1/(sigma.Y*sigma.Y)
tau.B <- 1/(sigma.B*sigma.B)
```

و مقایسه مقادیر به‌دست آمده از میانگین و انحراف استاندارد پسین به‌دست می‌آید. نتایج در این مثال برای چنین تغییراتی قوی و موثق هستند.

۹-۲-۳ مثال ۱- ب

۹-۲-۳-۱ اطلاعات در مورد پارامتر زمینه β به‌شکل توزیع احتمال ارائه می‌شود که به‌وسیله ارزشیابی نوع B از عدم‌قطعیت به‌دست آمده‌اند. در این مورد، مدل مشاهده تنها به شکل مجموعه‌ای از داده‌ها (i) در مثال (۱- الف) زیر بند (۹-۲-۲) می‌باشد که برابر است با:

$$Y_i \mid \theta\ ,\beta,\ \sigma_Y^2 \sim N\left(\theta+\beta,\sigma_Y^2\right).$$

۹-۲-۳-۲ اکنون سه پارامتر وجود دارند که به توزیع‌های پیشین اختصاص داده می‌شوند. برای پارامتر زمینه β، چگالی پیشین بر مبنای اطلاعات داده شده در مقدمه می‌باشد که عبارت است از:

$$\beta\ \sim\ \text{Uniform}(1{,}126{,}1{,}329).$$

برای θ و σ_Y،

$$\theta \sim \text{Uniform}(0,\ c)\ , \qquad \sigma_Y\ \sim\ \text{Uniform}(0,\ c)\ ,$$

که مقدارc، بزرگ است و ویژگی‌های توزیع‌های پیشین را کامل می‌کند.

۹-۲-۳-۳ کد WinBUGS برای این مثال به‌شکل زیر است:

```
Example1b{
theta~dunif(0,100)
beta~dunif(1.126,1.329)
sigma.Y~dunif(0,1)
gamma <- theta+beta
tau.Y <- 1/(sigma.Y*sigma.Y) for(i in 1:n){
y[i]~dnorm(gamma,tau.Y)
}
```

با داده‌های ارائه شده در مقدمه، این کد میانگین پسین برای θ ۲٫۳۰۹ و انحراف استاندارد پسین برای θ را برابر ۰٫۲۳۲ حاصل کرد. بازه اطمینان ٪۹۵ برای θ (۲٫۷۸۸ ۱٫۸۳۲) است. تحلیل حساسیت نتایج، دوباره راضی‌کننده و قابل قبول است.

۹-۲-۴ مثال ۱- پ

۹-۲-۴-۱ تنها تفاوت با مثال (۱- ب) مقادیر اندازه‌گیری شده واقعی است (که اکنون به زمینه نزدیک هستند) و بنابراین مدل مشابه و کد WinBUG می‌تواند مورد استفاده قرار گیرد. میانگین پسین برای θ اکنون ۰/۰۶۹ است و انحراف استاندارد پسین نیز برای θ برابر با ۰٫۰۶۷ و بازه اطمینان ٪۹۵ برای θ (۰٫۱۸۸ ۰٫۰۰۰) هستند. این نتایج برای تغییرات در مقدارc به اندازه کافی یکنواخت و قوی هستند. تغییرات شکل چگالی پیشین برای σ_Y از شکل یکنواخت به نتایج گاما در میانگین پسین: ۰/۰۵۸ و انحراف استاندارد پسین: ۰/۰۵۲ و بازه اطمینان ٪۹۵ (۰٫۱۵۰ ۰٫۰۰۰) است. این تغییرات، بزرگ‌تر از تغییرات مثال‌های قبلی است و نشان می‌دهد که به‌خاطر نزدیک بودن داده‌ها به زمینه، داده‌ها اطلاعات کاملی در مورد اندازه‌ده‌ها ندارند. اندازه σ_Y (تا حدی به‌وسیله توزیع پیشین کنترل می‌شود چون تنها پنج مقدار اندازه‌گیری شده وجود دارد که برآورد بر اساس آن است)، بر روی این که داده‌ها چقدر حاوی اطلاعات هستند اثر می‌گذارد. در یک مورد شبیه این راه‌حل محافظه‌کارانه، استفاده از یک بازه اطمینان طولانی‌تر مبتنی بر توزیع یکنواخت است. یک روش بهتر می‌تواند به‌دست آوردن مقادیر اندازه‌گیری شده بیشتر باشد. نتیجه نیز کاهش اثر چگالی پیشین σ_Y بر روی نتایج خواهد بود. (یک واقعیت جالب در مورد بازه‌های اطمینان بیزی مانند آنچه‌که در این‌جا داده شده است در منبع [۲۳] می‌تواند دیده شود. نویسنده‌ها نشان می‌دهند که در مدل‌هایی مانند مثال ۱، بازه اطمینان ٪۹۵ بیزی مبنی بر پوشش فراوانی‌گرای یکنواخت پیشین که به ٪۹۵ نزدیک است در حالی که بازه مبتنی برگامای پیشین معمولاً پوشش فراوانی‌گرای کمتری دارد.

۹-۲-۵ خلاصه مثال

مثال (۱- الف) وضعیت را زمانی‌که مقادیر اندازه‌گیری شده در دو منبع مستقل در یک ارزشیابی عدم‌قطعیت واحد مورد استفاده قرار می‌گیرد، را نشان می‌دهد. مثال (۱- ب) نشان می‌دهد که چگونه اطلاعات در مورد زمینه که برای اجرای ارزشیابی نوع B از عدم‌قطعیت استفاده می‌شود می‌تواند در مدل بیزی وارد شود. مثال (۱- پ) نشان می‌دهد یک محدودیت به‌راحتی می‌تواند در مدل بیزی وارد شود، نظیر محدودیت مثبت در این‌جا که بر روی ارزش اندازه‌ده وجود دارد. هم‌چنین نشان می‌دهد که چگونه انتخاب یک توزیع پیشین به‌طور غیرمستقیم می‌تواند بر روی نتایج اثر بگذارد.

۱۰ استنباط اتکایی برای ارزشیابی عدم‌قطعیت

۱-۱۰ روش پایه

۱-۱-۱۰ برای تابع اندازه‌گیری (۱) ارزشیابی عدم‌قطعیت برای اندازه‌ده θ ممکن است بر مبنای توزیع اتکایی برای θ باشد. مثال زیر دستورالعمل به‌دست آوردن توزیع‌های اتکایی برای پارامترهای موردنظر را نشان می‌دهد.

۲-۱-۱۰ فرض کنید $Y \sim N(\theta,1)$، که θ یک اندازه‌ده است.، فرآیند اندازه‌گیری واریانس معلوم برابر ۱، و Y متغیر تصادفی متناظر نشان دهنده مقادیری است که ممکن است مشاهده شوند. یک حالت نشان دهنده این است که ارتباط بین مقادیر اندازه‌گیری شده و فرآیند خطای آزمایش تصادفی برآمده از آن به‌وسیله معادله زیر نشان داده می‌شود:

$$Y = \theta + E, \tag{۱۹}$$

که E یک خطای تصادفی با توزیع $N(۰,۱)$ است. هر مقدار اندازه‌گیری شده با یک خطای آزمایشی تصادفی خاص ارتباط خاص دارد. فرض کنید که یک مقدار اندازه‌گیری شده ۱۰ به‌دست می‌آید. خطای اندازه‌گیری مرتبط با e نشان داده می‌شود. بنابراین:

$$10 = \theta + e$$

از این رو $\theta = 10 - e$ است. اگر مقدار e معلوم بود اندازه‌ده نیز باید دقیقاً معلوم باشد اما مقدار e معلوم نیست. با این‌حال، این واقعیت که توزیعی که e از آن به‌وجود می‌آید معلوم است کمک می‌کند که یک مجموعه از مقادیر θ تعریف شوند که این‌کار امکان‌پذیر است. برای مثال این‌که مقدار $\theta =$ ۲ برای اندازه‌ده تا چه حد امکان‌پذیر است؟ برای درستی این رابطه لازم است $e =$ ۸ باشد. مقدار ۸ به احتمال زیاد از توزیع $N(۰,۱)$ به‌دست می‌آید. بنابراین نتیجه می‌گیریم که مقدار $\theta =$ ۲ غیرمحتمل است. چقدر احتمال دارد که θ بین ۱۰ و ۱۲ قرار گیرد؟ برای θ بین ۱۰ و ۱۲، باید e بین صفر و ۲ باشد احتمال آن برابر با $\Phi(2) - \Phi(0)$ است که $\Phi(z)$ مقدار استاندارد تجمعی یا مضاعف توزیع گاوسی در z می‌باشد. بنابراین احتمال‌های مرتبط با E می‌تواند به احتمال‌های θ منتقل شوند. دانش در باره θ بر اساس مقدار اندازه‌گیری شده از ۱۰ می‌باشد و می‌تواند به‌وسیله توزیع متغیر تصادفی $\tilde{\theta}$ که توزیع آن به‌وسیله $10 - E$ داده شده است، توضیح داده شود که $\hat{\theta} \sim N(10,1)$ یا توزیع اتکایی برای θ (که توزیعی برای $\tilde{\theta}$ است) $N(۰,۱)$ است. متغیر تصادفی $\tilde{\theta}$ کمیت اتکایی (FQ)[1] برای θ نیز نامیده می‌شود. یک چنین FQ با چیزی که کمیت محوری کلی [۲۴]، [۲۵] یا کمیت محوری کلی اتکایی [۲۶]، [۲۷] نامیده می‌شود ارتباط دارد.

---

1 -Fiducail quantity

۱۰-۱-۳ در مثال بالا، فرض کنید دو اندازه‌گیری انجام گرفته است. اجازه بدهید Y_1 و Y_2 را متغیر تصادفی که ممکن است مقادیر ممکن را از دو نوع اندازه‌گیری به‌دست آیند در نظر بگیریم که می‌توان به‌صورت زیر نشان داد:

$$Y_1 = \theta + E_1,$$ (۲۰)

$$Y_2 = \theta + E_2.$$

فرض کنید مقادیر اندازه‌گیری شده ۱۰ و ۸ باشد. بنابراین، معادله زیر مقادیر اندازه‌گیری شده، اندازه‌ده و مقدار به‌دست آمده از خطای تصادفی را با e_1 و e_2 نشان می‌دهد.

$$۱۰ = \theta + e_1.$$

$$۸ = \theta + e_2.$$

مقادیر ممکن برای θ با مقادیر ممکن برای $\left(e_1 \text{ و } e_2\right)$ ارتباط دارد . چیزی که این مثال را با مثال قبلی متمایز می‌کند این است که در اینجا معلوم است $e_1 - e_2$ برابر ۲ است. بنابراین، یک مجموعه مقادیر ممکن برای الزام $\left(e_1 \text{ و } e_2\right)$ محدود شده است. معلوم است $\left(e_1 \text{ و } e_2\right)$ به توزیع گاوسی دو متغیره استاندارد تعلق دارد اما به‌خاطر قرار گرفتن بر روی خط $e_1 - e_2 = ۲$ محدود شده است. پس احتمالاتی که با θ ارتباط دارند احتمالاتی با $e_1 - 10$ یا $e_2 - 8$ هستند، $\left(e_1 \text{ و } e_2\right)$ یک مقدار معین از توزیع گاوسی استاندارد دو متغیره است به شرط این‌که یک موقعیت اضافی وجود داشته باشد که $e_1 - e_2 = ۲$ است. بنابراین یک FQیِ $\tilde{\theta}$ این‌گونه تعریف می‌شود: داشتن یک توزیع که برابر با توزیع وضعیتی از $E_1 - 10$ است $E_1 - E_2 = ۲$ می‌باشد. این یک توزیع مشابه با توزیع وضعیتی است که $E_2 - 8$ که به‌صورت $E_1 - E_2 = ۲$ داده شده است، می‌باشد. یک محاسبه ساده نشان می‌دهد که توزیع $\tilde{\theta}$ برابر با $N(\bar{y}, 1/2)$ است که $\bar{y} = \left(y_1 + y_2\right)/2 = (10 + 8)/2 = 9$.

۱۰-۱-۴ به‌صورت کلی‌تر، برای n اندازه‌گیری مستقل از $N(\theta,\sigma^2)$ به‌دست می‌آید.

$$
\begin{aligned}
Y_1 &= \theta + \sigma E_1,\\
Y_2 &= \theta + \sigma E_2,\\
&\dots,\\
Y_n &= \theta + \sigma E_n,
\end{aligned}
$$ (۲۱)

که در آن:

E_1،...،E_n متغیر تصادفی گاوسی استاندارد مستقل هستند.

توزیع اتکایی مشترک برای (θ,σ) می‌تواند به‌صورت زیر به‌دست آید. از دو n اول (یا هر دو n) بالا به‌عنوان معادلات ساختاری برای حل θ و σ که به‌وسیله $\tilde{\theta}$ و $\tilde{\sigma}$ نشان داده می‌شوند، مانند توابع y_1، y_2، E_1 و E_2

استفاده کنید. توزیع اتکایی مشترک برای (θ,σ) توزیع مشترک برای $(\tilde{\theta},\tilde{\sigma})$ است با این شرط که E_i در بقیه معادلات ۲- n نیز صدق کند. در حالت خاص، توزیع اتکایی برای θ برابر است با:

$$\tilde{\theta} = \overline{y} - \frac{s}{\sqrt{n}} T_{n-1}, \tag{۲۲}$$

برای مثال، یک توزیع t تغییر یافته و درجه بندی شده با درجه آزادی ۱- n را می‌توان نام برد. در این‌جا $\overline{y}$ و s مقادیر معینی از میانگین نمونه $\bar{X}$ و انحراف استاندارد نمونه S برای n مقدار اندازه‌گیری هستند و T_{n-1} یک متغیر تصادفی است که دارای یک توزیع t با ۱- n درجه آزادی می‌باشد.

۱۰-۱-۵ در این‌جا یک روش ساده‌تر از آنچه که برای به‌دست آوردن توزیع اتکایی برای θ در فرمول (۲۲) استفاده شد وجود دارد که در مثال‌های بعدی نشان داده خواهد شد.

۱۰-۱-۶ بحث فوق می‌تواند حالت کلی‌تر به‌خود بگیرد و توزیع‌های اتکایی می‌توانند برای پارامترهای مدل در مسائل با گستره وسیع‌تر توسعه یابد. نقطه آغازین برای این فرآیند معادله ساختاری نامیده می‌شود [۲۸]. که با $Y = G(\beta,E)$ نشان داده می‌شود. برای یک اندازه‌گیری واحد، معادله (۱۹) معادله ساختاری را تشکیل می‌دهد. برای n اندازه‌گیری، معادلات (۲۱) معادله‌های ساختاری را تشکیل می‌دهند. معادلات ساختاری ارتباط بین اندازه‌گیری‌های Y با پارامترهای مدل β و فرآیندهای خطای E که توزیع آن‌ها کاملاً معلوم است، هستند. برای مثال برای یک اندازه‌گیری واحد، توزیع برای E کاملاً معلوم است. برای هر مقدار ثابت β، توزیع برای E و معادله ساختاری $G(.)$، توزیع برای داده‌های Y را تعیین می‌کند. بعد از مشاهده داده‌های Y نقش داده‌ها و پارامترها می‌تواند عوض شود. در حالت خاص، مقدار Y ثابت است و توزیع E و معادله ساختاری $G(.)$ برای حدس و استنباط یک توزیع برای β مورد استفاده قرار می‌گیرد. این همان چیزی است که مبحث اتکایی را تشکیل می‌دهد.

۱۰-۲ مثال ۱

۱۰-۲-۱ مثال ۱- الف

۱۰-۲-۱-۱ برای توضیح، مثال (۱- الف) که در بند ۷ که در آن کمیت فیزیکی θ از مقادیر اندازه‌گیری شده برآورد می‌شود را بررسی می‌کنیم که از مدل زیر پیروی می‌کند:

$$Y_i = \theta + \beta + \varepsilon_i, \qquad i = 1,....,n \tag{۲۳}$$

که ε_i خطای اندازه‌گیری مستقل با $\varepsilon_i \sim N\left(0,\sigma_Y^2\right)$ است. هم‌چنین β نشانگر زمینه بوده و می‌تواند از اندازه‌گیری‌هایی که از مدل پیروی می‌کنند برآورد شود.

$$B_i = \beta + \delta_i \qquad i = 1,...,n_b, \tag{۲۴}$$

که δ_i خطاهای اندازه‌گیری مستقل با $\delta_i \sim N\left(0,\sigma_B^2\right)$ است. فرض می‌شود ε_i و δ_i مستقل هستند. از فرمول (۲۳) و (۲۴) نتیجه می‌شود که $\overline{Y}-\overline{B}$ یک توزیع گاوسی با میانگین θ و واریانس $\sigma_Y^2/n+\sigma_B^2/n_b$ که $\overline{Y}$ و $\overline{B}$ میانگین Y_i و B_i هستند و می‌توانند به‌صورت زیر بیان شوند:

$$\overline{Y}-\overline{B}=\theta+\sqrt{\frac{\sigma_Y^2}{n}+\frac{\sigma_B^2}{n_b}}Z, \tag{۲۵}$$

که در آن:

Z متغیر تصادفی گاوسی استاندارد است. این یک معادله ساختاری برای $\overline{Y}-\overline{B}$ است. هم‌چنین

$$W_y=\frac{(n-1)S_y^2}{\sigma_Y^2}\sim\chi^2(n-1)$$

و

$$W_b=\frac{(n_b-1)S_b^2}{\sigma_B^2}\sim\chi^2(n_b-1),$$

که $\chi^2(\nu)$ نشانگر توزیع کای‌دو با درجه آزادی ν است. S_y^2 و S_b^2 واریانس نمونه Y_i و B_i هستند. بنابراین:

$$S_y^2=\frac{\sigma_Y^2 W_y}{n-1} \tag{۲۶}$$

یک معادله ساختاری برای S_y^2 و

$$S_b^2=\frac{\sigma_B^2 W_b}{n_b-1} \tag{۲۷}$$

یک معادله ساختاری برای S_b^2 است. با حل سه معادله ساختاری بالا برای θ، σ_Y و σ_B، یک FQ برای θ به‌صورت زیر به‌دست می‌آید:

$$\tilde{\theta}=\overline{y}-\overline{b}-\sqrt{\frac{(n-1)S_y^2}{nW_y}+\frac{(n_b-1)S_b^2}{n_bW_b}}Z. \tag{۲۸}$$

۱۰-۲-۱-۲ یک بازه اتکایی α - ۱ برای θ به‌وسیله $(\tilde{\theta}_{\alpha/2} \quad \tilde{\theta}_{1-\alpha/2})$ داده می‌شود کـه $\tilde{\theta}_\alpha$ همـان چنـدک α اُم توزیع $\tilde{\theta}$ است. این چندک‌ها می‌توانند به صورت تحلیلی در وضعیت‌های ساده تعیین شـوند. بـه هرحـال، بـه صورت ساده‌تر با استفاده از رویکرد مونت کارلو تقریب زده می‌شوند. این رویکرد شامل به‌دست آوردن تعـداد بسیار زیادی از مفاهیم و درک آنها از توزیع $\tilde{\theta}$ و تعیین چندک‌های $\alpha/2$ و $1-\alpha/2$ به صورت تجربی است. ایـن چندک‌ها برای برآورد $\tilde{\theta}_{\alpha/2}$ و $\tilde{\theta}_{1-\alpha/2}$ مورد استفاده قرار می‌گیرند. یک مقدار واحد از $\tilde{\theta}$ ممکن است بـه‌صـورت زیر به‌وجود آید.

الف- متغیر تصادفی گاوسی استاندارد Z را تعیین و به‌وجود آورید.

ب- مقادیر مستقل χ^2 متغیر تصادفی W_b و W_y را با درجه آزادی n-۱ و n_b-۱ به‌وجود بیاورید.

پ- $\tilde{\theta}$ را در فرمول (۲۸) محاسبه کنید.

برای ایـن مثـال، $n = n_b = 5, \ \bar{y} = 3{,}537, \ s_y = 0{,}342, \ \bar{b} = 1{,}228,$ و $s_b = 0{,}131$ اسـت. یـک برنامه R برای به‌وجود آوردن ۵۰۰۰۰۰ مقدار $\tilde{\theta}$ به صورت زیر است.

```
nrun = 500000
Z = rnorm(nrun)
W1 = rchisq(nrun, 4)
Wb = rchisq(nrun, 4)
theta = 3.537 - 1.228 - sqrt(4*0.342^2/(5*W1)+4*0.131^2/(5*Wb))*Z
```

میانگین توزیع شبیه سازی شده

```
mean(theta)
## 2.308893
```

است و یک بازه اتکایی ٪۹۵ بر اساس چندک‌های ۰/۰۲۵ و ۰/۹۷۵ توزیع شبیه‌سازی شده برابر است با:

```
quantile(theta, c(0.025,0.975))
## 2.5 % 97.5 %
## 1.857814 2.760931
```

به‌عنوان مثال، بازه اتکایی ٪۹۵ به‌وسیله (۱/۸۵۸ ۲/۷۶۱) داده می‌شود.

۱۰-۲-۲ مثال ۱- ب

۱۰-۲-۲-۱ حالا هیچ داده آماری مرتبط با زمینه وجود ندارد. فرض می‌شود که اطلاعات مربوط به β در مواردی از توزیع احتمال برای β مشخص شده است و β و ε_i نیز مستقل هستند. علاوه بر این فرض می‌شود که توزیع احتمال برای β کاملاً معلوم است، و بنابراین هیچ‌گونه پارامتر نامعلومی در آن وجود ندارد.

۱۰-۲-۲-۲ معادله ساختاری برای $\bar{Y}$ به‌وسیله

$$\bar{Y} = \theta + \beta + \frac{\sigma}{\sqrt{n}} Z. \qquad (۲۹)$$

همراه با معادله ساختاری برای S_y^2 در فرمول (۲۶)، که یک FQ برای θ از

$$\tilde{\theta} = \bar{y} - \beta - \frac{s_y}{\sqrt{n}} - \frac{Z}{\sqrt{W_y/(n-1)}}.$$

به‌دست می‌آید از این رو $Z/\sqrt{W_y/(n-1)} = T_{n-1}$ یک متغیر تصادفی و دارای توزیع t با درجه آزادی n -۱ است.

$$\tilde{\theta} = \bar{y} - \beta - \frac{s_y}{\sqrt{n}} T_{n-1} \qquad (۳۰)$$

یک مقدار واحد برای $\tilde{\theta}$ ممکن است به‌صورت زیر به‌دست آید.

الف- یک مقدار T_{n-1} از متغیر تصادفی t- استیودنت با درجه آزادی ۱- n را به‌دست آورید.

ب- β را با توجه به توزیع آن و مستقل از T_{n-1} به‌دست آورید.

پ- $\tilde{\theta}$ را در فرمول (۳۰) محاسبه کنید.

برای این مثال، فرض می‌شود، β به‌طور یکنواخت در بازه (۱٫۱۲۶ ۱٫۳۲۹) توزیع می‌شود. برای $\tilde{\theta}$ ، ۵۰۰۰۰۰ مقدار به‌صورت زیر به‌دست می‌آید:

```
beta = runif(nrun, 1.126, 1.329)
theta = 3.537 - beta - 0.342/sqrt(5)*rt(nrun, 4)
```

میانگین توزیع شبیه‌سازی شده است

```
mean(theta)
## 2.309454
```

و بازه اتکایی ٪۹۵ بر روی چندک‌های ۰٫۰۲۵ و ۰٫۰۹۷۵ توزیع شبیه‌سازی شده است.

```
quantile(theta, c(0.025, 0.975))
## 2.5 % 97.5 %
## 1.871685 2.745590
```

به‌طور مثال بازه اتکایی ٪۹۵ به‌وسیله (۲٫۷۴۶ ۱٫۸۷۲) داده می‌شود.

۱۰-۲-۲-۳ بازه اتکایی فوق با بازه عدم‌قطعیت به‌دست آمده بـا اسـتفاده از روش پیشـنهادی در GUMS1 مطابقت دارد.

۱۰-۲-۳ مثال ۱- پ

۱۰-۲-۳-۱ مورد مثال (۱- ب) به جز $\bar{y}$ = ۱٫۱۹۶ و s_y = ۰٫۱۰۶ به‌کار مـی‌رود. ۵۰۰۰۰۰ مقـدار بـرای $\tilde{\theta}$ از طریق

```
theta = 1.196 - beta - 0.106/sqrt(5)*rt(nrun, 4)
```

به‌دست می‌آید. میانگین آن:

```
mean(theta)
## -0.03158058
```

است. مقدار مقادیر خارج از فضای پارامتر می‌تواند با اسـتفاده از θ باشد که خارج از فضای پارامتر برای

```
length((1:nrun)[theta < 0])
## 319168
```

به‌دست آید.

یک رویکرد برای بررسی محدودیت‌های پارامتر، کوتاه کردن توزیع اتکایی در فضای پارامتر محدود است. بـه این معنا که، از $\max(\tilde{\theta},0)$ برای به‌دست آوردن مقادیر توزیع اتکایی برای θ اسـتفاده مـی‌کنـیم. یـک بـازه اتکایی ٪۹۵ به‌شکل زیر محاسبه می‌شود:

```
quantile(pmax(theta, 0), c(0.025, 0.975))
## 2.5 % 97.5 %
## 0.0000000 0.1361553
```

به‌عنوان مثال، بازه اتکایی ٪۹۵ به‌وسیله (۰٫۱۳۶ ۰٫۰۰۰) داده می‌شود.

۱۰-۲-۳-۲ دستورهای توصیف شده در زیربندهای ۱۰-۲-۱-۱ و ۱۰-۲-۲-۲ می‌توانند به مدل‌های آماری قراردادی تعمیم داده شوند. یک دستور برای ایجاد FQs در منبع [۲۹] داده شده است. یک دستور ساده‌تر برای مسائل رایج‌ترکه در آماره‌های بسنده وجود دارند در استاندارد (منبع [۳۰]) داده شده است و علاوه بـر این در منبع [۲۴] و [۲۵] بحث خواهند شد. که در اینجا به‌علت کامل شدن، دوباره گفته شدند. این دستور شامل گام‌های زیر است:

الف- هر آماره بسنده را مثل یک تابع از یک یا چند پارامتر و متغیر تصادفی که توزیـع آن‌هـا کـاملاً معلـوم است و هیچ پارامتر نامعلوم و ناشناخته‌ای وجود ندارد بیان کنید. سپس یک معادله ساختاری برای هر آماره بسنده به‌دست بیاورید.

ب- در هر معادله ساختاری هر پارامتر را مانند یـک تـابع از آمـاره‌هـای بسـنده و متغیرهـای تصـادفی کـه توزیع‌آن‌ها کاملاً معلوم است بیان کنید.

پ) یک FQ برای هر پارامتر با جای‌گذاری آماره‌های بسنده با مقادیر مشاهده شده متناظر به‌دست آورید.

۱۱ مثال۲: کالیبراسیون یک بلوک سنجه

۱-۱۱ کلیات

۱-۱-۱۱ این مثال که از پیوست H.1 از استاندارد GUM گرفته شده است با تعریف و تعیین طول یک بلوک سنجه و مقایسه آن با بلوک سنجه مشابه دیگر که قبلاً کالیبره شده است بیان می‌شود. نماد مورد استفاده در استاندارد GUM در اینجا دنبال می‌شود اما هر جا که نیاز باشد اصلاح خواهد شد تا با علائم قراردادی گفته شده در بند ۴ که اندازه‌ده‌های اندازه‌گیری را از مقادیر اندازه‌گیری شده متمایز می‌کند، همخوانی داشته باشد. جدول ۲ کمیت‌های فیزیکی استفاده شده را نشان می‌دهد.

۲-۱-۱۱ از نمادهای داده شده در جدول ۳ استفاده کنید که مبتنی بر:

- خط اول معادله (H.2) در استاندارد GUM،
- روابط $\alpha = \alpha_s + \delta_\alpha$ و $\theta_s = \overline{\theta} + \Delta - \delta_\theta$ تعریف شده در H.1.2 و
- استنباط نشر عدم‌قطعیت در H.1.3.2 و H.1.3.4 ، از استاندارد GUM گرفته شده است.

۳-۱-۱۱ مدل اندازه‌گیری برای λ مورد استفاده دراستاندارد GUM و تحلیل مثال H.1 می‌تواند به‌صورت زیر بیان شود:

$$\lambda = \frac{\lambda_s\left[1 + \alpha_s\left(\overline{\theta} + \Delta - \delta_\theta\right)\right] + S_\lambda + S_{C_r} + S_{C_{nr}}}{1 + (\alpha_s + \delta_s)\left(\overline{\theta} + \Delta\right)} \qquad (۳۱)$$

معادله (۳۱) مدل اندازه‌گیری است که در H-1-1 توضیح داده شد و هم‌چنین این معادله خط اول معادله(H.2) از استاندارد GUM است، صرف نظر از تقریب به‌دست آمده در خط دوم معادله (H.2) و بنابراین کاربرد آن در بقیه معادله H.1 است.

جدول۳- نماد سازی مربوط به تحلیل مثال H.1.1، استاندارد GUM تحت سه رویکـرد آمـاری. متغیـر تصـادفی متناظر با δ_λ با $\bar{D}_\lambda$ و مقدار به‌دست آمده آن با $\bar{d}_\lambda$ نشان داده می‌شود.

کمیت	نماد
طول سنجه نامعلوم در ۲۰ °C	λ
طول سنجه استاندارد در ۲۰ °C	λ_s
اختلاف بین طول سنجه در دمای اتاق آزمایشگاه	δ_λ
تصحیح تفاوت بین طول بلوک سنجه به منظور جبران خطای مقایسه‌گر تصادفی	δ_{c_r}
تصحیح تفاوت بین طول‌های بلوک سنجه جهت مقایسه با خطاهای مقایسه‌گر نظام‌مند	$\delta_{c_{nr}}$
ضریب انبساط دمایی نهایی بلوک استاندارد	α_s
تفاوت در ضرایب انبساط دمایی بلوک‌های استاندارد و نامعلوم	δ_α
انحراف میانگین دمای بستر آزمایش از شرایط استاندارد در طول جمع‌آوری داده‌ها	$\bar{\theta}$
اختلاف چرخشی دمای بستر آزمایش از دمای میانگین به‌دلیل کنترل دماپا[1]	Δ
تفاوت در دمای سنجه‌های استاندارد و نامعلوم	δ_θ

1 - Thermostatic

۱۱-۱-۴ معادله (۳۱) در مورد کمیت‌های فیزیکی جهت تعیین طول بلوک سنجه نیز بیان شد که پیشتر برای خلاصه نمودن اثرات به‌دلیل تفاوت بین طول‌های دو بلوک سنجه در دمای آزمایش خواب نیز مورد استفاده قرار گرفته بود. این یک تمرین خوب برای بیان مدل اندازه‌گیری در مورد تمام کمیت‌های مورد نیاز برای تعیین آن است. این تمرین کمک می‌کند تا عدم موفقیت احتمالی برای تعیین کالیبراسیون برای تعیین ارتباط بین کمیت‌های فیزیکی مختلف مانند θ و θ_s و α و α_s که در H.1.2 ذکر شد به حداقل برسد که مقادیر آن در خاتمه احتمالات بر اساس داده‌های مشابه است.

۱۱-۱-۵ جدول ۴ بقیه اطلاعات به‌دست‌آمده از تحلیل مثالH.1، استاندارد GUM را که برای تحلیل رویکردهای آماری متفاوت مورد نیاز است به‌طور خلاصه نشان می‌دهد که در ادامه، بحث و مقایسه خواهند شد.

۱۱-۱-۶ توصیف مثال در استاندارد GUM نشان می‌دهد که تنها یک کمیت، δ_l وجود دارد که مقدار آن با استفاده از تحلیل داده‌های آماری برآورد شده است. توزیع میانگین مقادیر اندازه‌گیری شده، که یک برآورد از δ_l را فراهم می‌کند به‌عنوان یک مقدار گاوسی (نرمال) با یک مقدار موردانتظار که بستگی به طول بلوک سنجه با یک مقدار مورد انتظار و سایر کمیت‌های فیزیکی گفته شده در جدول ۳ و ۴ توصیف شده، به‌دست می‌آید.

۱۱-۱-۷ مقادیر و عدم‌قطعیت استانداردهای مرتبط با برآورد سایر کمیت‌ها به‌وسیله ارزشیابی‌های نوع B به‌دست می‌آیند. چون δ_α و δ_θ از توزیع مستطیلی تبعیت می‌کند تا توزیع گاوسی، اما به هرحال هیچ‌گونه روش آماری مورد قبول برای محاسبه درجه آزادی در این دو مورد وجود ندارد. در نتیجه درجه‌های آزادی داده شده برای آن کمیت‌ها مورد استفاده قرار خواهد گرفت.

جدول۴- خلاصه‌ای از تحلیل H_1 از مثال استاندارد GUM که برای تحلیل دوباره نیاز است.

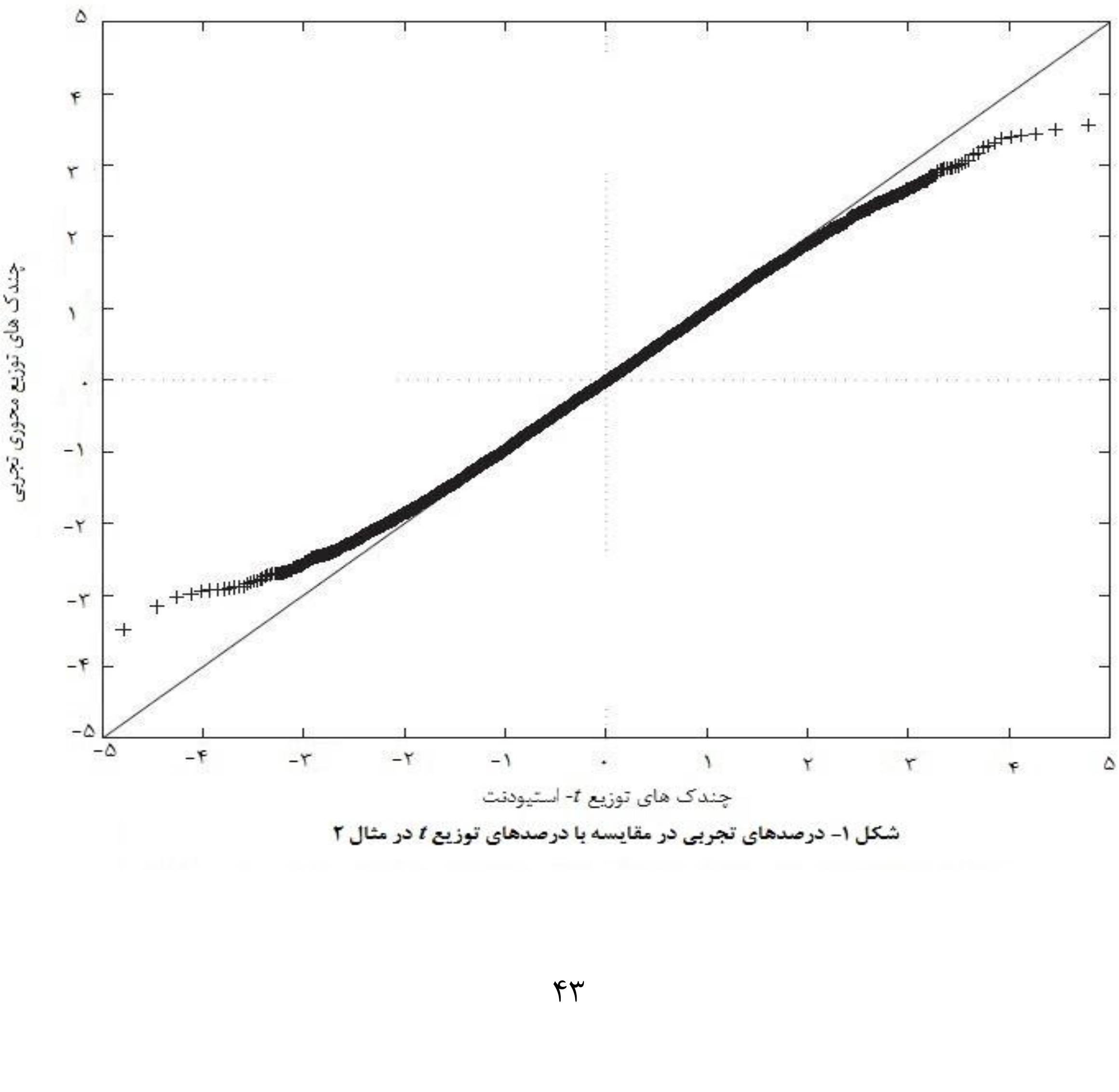
کمیت	مقدار	عدم‌قطعیت استاندارد	درجه‌های آزادی	نوع ارزشیابی عدم‌قطعیت	توزیع مشخص کننده
λ_s	۵۰۰۰۰۶۲۳ nm[1]	۲۵ nm	۱۸	B	گاوسی
$\bar{d}_\lambda$	۲۱۵ nm	۵/۸ nm	۲۴	A	گاوسی
δ_{c_r}	۰ nm	۳/۹ nm	۵	B	گاوسی
$\delta_{c_{nr}}$	۰ nm	۶/۷ nm	۸	B	گاوسی
α_s	۱۱/۵ × ۱۰^{-۶} °C^{-۱}	۱/۲ × ۱۰^{-۶} °C^{-۱}		B	مستطیلی
δ_α	°C^{-۱}	۰/۵۸ × ۱۰^{-۶} °C^{-۱}	۵۰	B	مستطیلی
$\bar{\theta}$	-۰/۱°C	۰/۲ °C		B	مشخص نیست
Δ	۰ °C	۰/۳۵ °C		B	معکوس سینوس
δ_θ	۰ °C	۰/۰۲۹ ° C	۲	B	مستطیلی

[1] - نانومتر

۱۱-۲ رویکرد فروانی‌گرا

۱۱-۲-۱ در این مثال، ضرایب حساسیت $C_{\alpha_s} = C_{\theta_S}$ از بین می‌روند و عبارت‌های دستوری دوم در معادلات ۶ و ۷ وارد می‌شوند اگر چه فقط یکی از آن‌ها به‌طور قابل ملاحظه‌ای متفاوت از صفر است (GUM, p 71) .

۱۱-۲-۲ مقدار y = ۵۰۰۰۰ ۸۳۸ nm تعداد از اندازه‌ده‌ها، به‌طور مثال، طول بلوک سنجه تحت کالیبراسیون و عدم‌قطعیت استاندارد مرتبط $u(y)$ = ۳۴ nm با ارزشیابی در استاندارد GUM با استفاده از موارد دستوری دوم به‌دست آمد. همان‌طوری‌که در زیربند ۸-۱-۱۲ ذکر شد عدم‌قطعیت‌های مرتبط با این برآوردها با خطای درجه‌دوم حاشیه‌ای به‌طور تقریبی پارامترهای λ_s و θ_s بر روی توزیع‌های نرمال آن‌ها به‌طور میانگین در نظر گرفته شدند و $\theta_s \lambda_s$ و δ_θ با توجه به توزیع‌های یکنواخت آن‌ها که تلفیق می‌شوند به‌دست می‌آیند. این نتایج با نشر توزیع‌هایی که با روش مونت‌کارلو در GUMS1 اجرا شدند، تأیید شدند که جواب خیلی نزدیکی به‌دست آمد. علاوه بر این، تقریب به وسیله توزیع t در فرمول (۱۰) به‌نظر می‌رسد منطقی باشد. شکل ۱ درصدهای تجربی که در مقایسه با کمیت‌های توزیع t نشان داده شده‌اند را ارائه می‌دهد که در آن درجه‌های آزادی با توجه به معادله (۸) برآورد شده‌اند. نتایج بیشتر از روش مونت‌کارلو در منبع [۳۱] گزارش شده‌اند.



شکل ۱- درصدهای تجربی در مقایسه با درصدهای توزیع t در مثال ۲

۱۱-۲-۳ برای به‌دست آوردن یک بازه خودگردان برای این مثال، λ : ۵۰۰۰۰۸۳۸ nm برآورد شد و عدم‌قطعیت استاندارد ترکیبی آن $u = 31{,}7$ nm است. از معادله (۱۳) بازه اطمینان٪ ۱۰۰(۱ - α) –t خودگردان $\left(50000838 - \hat{t}_{1-\alpha/2}\cdot 31{,}7 \quad 50000838 + \hat{t}_{\alpha/2}\cdot 31{,}7\right)$ است که $\hat{t}_\beta$، ۱۰۰β اُمین درصد W^* از معادله (۱۲) است کد R برای به‌دست آوردن B = ۱۰۰۰۰ مقدار W^* به شرح زیر است:

```
B = 10000
x.star = cbind(
rnorm(B, mean=50000623, sd=25),
rnorm(B, mean=215, sd=5.8),
rnorm(B, mean=0, sd=3.9),
rnorm(B, mean=0, sd=6.7),
runif(B, min=0.0000095, max=0.0000135),
runif(B, min=-0.000001, max=0.000001),
runif(B, min=-0.45, max=0.25),
rbeta(B, 0.5, 0.5)-0.5,
runif(B, min=-0.05, max=0.05))
u.star = cbind(
25 * sqrt(rchisq(B, df=18)/18),
5.8 * sqrt(rchisq(B, df=24)/24),
3.9 * sqrt(rchisq(B, df=5)/5),
6.7 * sqrt(rchisq(B, df=8)/8),
0.000001 2,
0.00000058 * sqrt(rchisq(B, df=50)/50),
0.2,
0.35,
0.029 * sqrt(rchisq(B, df=2)/2))
x.name = c("L.s","D.lambda","Dc.r","Dc.s","A.std","D.alpha",
"T.bar","T.cv","D.theta")
f = expression((L.s*(1+A.std*(T.bar+T.cv-D.theta))+
D.lambda+Dc.r+Dc.s)/ (1+(A.std+D.alpha)*(T.bar+T.cv)))
star = delta(f, x.star, u.star, x.name)
w.star = (star$y-50000838)/star$uc
```

تابع دلتای R به شرح زیر تعریف می‌شود.

```
delta = function(meq,x,u,namevec){
```

```
for(i in 1:ncol(x)) assign(namevec[i], x[,i])
c = attr(eval(deriv(meq,namevec)),"gradient")
list(y=eval(meq),uc=sqrt(apply((c*u)^2,1,sum)))
}
```

این تابع یک R حالت با معنی meq، تابع اندازه‌گیری که اسم پارامترهای آن باnamevec داده شده است و ماتریس مقادیر ورودی **x**، یک ستون از **x** شامل تکرارهایی از خودگردان برای هر کمیت در meq است. تابع از تابعR، derive برای ارزشیابی تابع اندازه‌گیری و به‌دست آوردن اولین مشتق‌های c (گرادیان تابع چگالی احتمال بر محور X) استفاده می‌کند که تمام پارامترها در namevec و با مقادیر ورودی x ارزشیابی شده‌اند. در خاتمه تابع حالت ارزشیابی شده (به‌عنوانy) را به‌دست می‌دهد و عدم‌قطعیت مرتبط را با استفاده از تقریب تیلور اولیه عادی محاسبه می‌کند. در کد خودگردان فوق تابع دلتا برای اندازه‌گیری تابع تعریف شده به‌عنوان f در کد به‌کار می‌رود.

بازه اطمینان t- خودگردان ٪۹۵ بر حسب چندک‌های ۰/۰۲۵ و ۰/۹۷۵ از توزیع شبیه‌سازی شده به شرح زیر است:

```
50000838-quantile(w.star,c(0.975,0.025))*31.70511
## 50000777 50000899
```

این بازه (۵۰۰۰۰۷۷۷۵۰۰۰۰۸۹۹) بر حسب nm ، تقریبا ٪۱۰ از بازه‌ای که از استاندارد GUM به‌دست می‌آید، کوتاه‌تر است. این رفتار کلی کوتاه‌تر بودن هر بازه خودگردان از بازه به‌دست آمده از یک ارزشیابی عدم‌قطعیت براساس تقریب مرتبه اول تیلور می‌باشد که به‌صورت بیشتر در منبع [۳۲] بحث شده است.

۱۱-۳ رویکرد بیزی

۱۱-۳-۱ مدل مشاهده‌ای که داده‌ها را به پارامترها مرتبط می‌کند می‌تواند، مشخص شود.

۱۱-۳-۲ استاندارد GUM می‌تواند به‌عنوان حالتی که مقدار مورد انتظار $E(D_\lambda)$ را از اندازه‌گیری داریم و برابر δ_λ است، تفسیر شود. که:

$$\delta_\lambda = \lambda\left(1 + (\delta_\alpha + \alpha_s)(\overline{\theta} + \Delta)\right) - \lambda_s\left[1 + \left[(\overline{\theta} + \Delta) - \delta_\theta\right]\alpha_s\right]. \quad (۳۲)$$

مقدار مورد انتظار از اندازه‌گیری‌های یک تابع از پارامتر بردار $\gamma = \lambda_s, \overline{\theta}, \Delta, \delta_a, \alpha_s, \delta_\theta$ و اندازه‌ده λ است. استاندارد GUM دو قسمت اضافی از عدم‌قطعیت که شامل مقایسه مشخص شده به‌وسیله ارزشیابی نوع B است را ارائه می‌دهد. بنابراین در مقدار مورد انتظار از تفاوت اندازه‌گیری که برابر با δ_λ است، عدم‌قطعیت وجود دارد. به‌همین ترتیب و مشابه استاندارد GUM، دو جزئی که می‌توانند با همدیگر به‌صورت اضافی ترکیب شده و یک عدم‌قطعیت برابر ۷/۸ nm با ۱۲درجه آزادی و با استفاده از فرمول ولش ساترثویت به‌دست بیاورند. مدل آماری دو مرحله‌ای زیر اطلاعات موجود را ترکیب می‌کند:

$$\bar{D}_\lambda \mid \delta_{\lambda_r} \sim N\left(\delta_{\lambda_r}, \sigma_{D_\lambda}^2/5\right)$$
$$\delta_{\lambda_r} \middle| \delta_\lambda \sim \delta_\lambda + 7.8 \cdot T_{12}. \qquad (۳۳)$$

۱۱-۳-۳ عدم‌قطعیت داده شده در مثال با تفاوت اندازه‌گیری S_{d_λ} مرتبط است و با استفاده از ارزشیابی نوع A به‌دست آمده است و یک برآورد از σ_{D_λ} را ارائه می‌دهد. با توجه به تئوری احتمال اصلی برای یک نمونه بـه اندازه n از توزیع گاوسی با واریانس معلوم σ^2 داریم:

$$\frac{(n-1)}{\sigma^2} S^2 \sim \chi^2(n-1)$$

و از آنجایی که $\chi^2(n-1)$ چگالی $\left(\frac{n-1}{2}, \frac{1}{2}\right)$ Gamma نیز است،

$$S_{d\lambda}^2 \mid \sigma_{D_\lambda} \sim Gamma\left[\frac{(25-1)}{2}, \frac{1}{2\sigma_{D_\lambda}^2}\right]$$

۱۱-۳-۴ در مدل آماری، ۸ پارامتر وجود دارد که شامل اندازه‌ده λ هستند. برای پیدا کـردن توزیـع پسـین برای λ، توزیع پیشین مشترک از ۸ پارامتر باید ابتدا مشخص شـوند. بـه‌طـور پـیش‌فـرض، ایـن متغیرهـای تصادفی می‌توانند مستقل در نظر گرفته شوند و بنابراین توزیع مشترک آن‌ها حاصل توزیع پیشین تـک‌تـک آن‌ها خواهد بود. برای مؤلفه‌های بردار پارامتر γ، اطلاعاتی که برای اجزای ارزشـیابی نـوع B از عـدم‌قطعیـت استفاده می‌شود می‌توانند به‌عنوان چگالی‌های پیشین که حاوی اطلاعات مفیدی هستند تفسیر شوند.

$$\lambda_s \sim N(50000623, 625), \qquad (۳۴)$$

$$\delta_\alpha \sim Uniform\left[-1\times10^{-6}, 1\times10^{-6}\right],$$
$$\bar{\theta} \sim N(-0{,}1, \quad 0{,}1681),$$
$$\Delta \sim Beta(0{,}5, \quad 0{,}05)-0{,}5,$$
$$\alpha_s \sim Uniform\left[9{,}5\times10^{-6}, \quad 13{,}5\times10^{-6}\right],$$
$$\delta_\theta \sim Uniform(-0{,}05, \quad 0{,}05),$$

بنابراین توزیع پیشین مشترک به‌صورت زیر است:

$$p(\gamma) = p(\lambda_s)\, p(\delta_\alpha)\, p(\bar{\theta})\, p(\Delta)\, p(\alpha_s)\, p(\delta_\theta).$$

توزیع‌های پیشین برای اندازه‌ده λ و σ_{D_λ} برای تکمیل ویژگی پیشین لازم هستند. در ایـن مثـال، هـیچ‌گونـه اطلاعات اضافی در مورد این دو پارامتر وجود ندارد به جز اینکه یکی از آنها غیرمنفی هستند. همانند مثال ۱ پارامترها به‌عنوان منبع اولیه اختصاص داده می‌شوند [۲۰]. برای λ یک چگالی منبع به‌صـورت تقریـب زیـر است:

$$\lambda \sim Uniform(0, c), \qquad (۳۵)$$

که یک مقدار بزرگ برای C در نظر گرفته می‌شود. همین‌طور برای σ_{D_λ}،

$$\sigma_{D_\lambda} \sim, Uniform(0,c) \quad (۳۶)$$

یا $\mathrm{Gamma}(c,c)$. که این ویژگی توزیع پیشین را کامل می‌کند.

توجه کنید که دو توزیع پیشین منبع که در آن تحلیل حساسیت یک اثر کوچک را بر روی نتایج نشان می‌دهد، تنها توزیع‌هایی هستند که در برخی موارد در رویکردهای اتکایی و فراوانی‌گرا مورد استفاده قرار نمی‌گیرند.

۱۱-۳-۵ کاربرد قضیه بیزی و نتایج آن در چگالی پسین مشترک $[\lambda, \gamma, \sigma_{D_\lambda}]$ به‌صورت زیر است:

$$p(\lambda, \gamma, \sigma_{D_\lambda} \mid \bar{d}_\lambda, s_{d_\lambda}) = \frac{p(\bar{d}_\lambda \mid \delta_{\lambda_r})\, p(\delta_{\lambda_r} \mid \delta_\lambda) p(\gamma) p(\lambda) p(\sigma_{D_\lambda})}{\int p(\bar{d}_\lambda \mid \delta_{\lambda_r})\, p(\delta_{\lambda_r} \mid \delta_\lambda) p(\gamma) p(\lambda) d\gamma\, d\lambda d\sigma_{D_\lambda}}$$

چگالی پسین λ بنابراین به‌وسیله انتگرال به‌صورت زیر به‌دست می‌آید:

$$p(\lambda \mid \bar{d}_\lambda, s_{d_\lambda}) = \int p(\lambda, \gamma, \sigma_{D_\lambda} \mid \bar{d}_\lambda, s_{d_\lambda}) d\gamma d\sigma_{D_\lambda}.$$

این توزیع پسین همه اطلاعات موجود مربوط به λ را بعد از اندازه‌گیری‌های به‌دست آمده خلاصه می‌کند.
کد WinBUGS برای این مثال به‌صورت زیر است:

```
Example2 {
n<-25
df<-(n-1)/2
lambda~dnorm(0.1,0E-18)
delta.a~dunif(-0.000001, 0.000001)
alpha~dunif(0.0000095,0.0000135)
theta~dnorm(-0.1,5.94)
ddelt~dbeta(0.5,0.5)
delta<-ddelt-0.5
delta.t~dunif(-0.05,0.05)
lambda.s~dnorm(50000623, 0.0016)
sigma.D~dunif(0.20)
tau.D<-1/(sigma.D*sigma.D)
delta.l<-lambda*(1+(delta.a+alpha)*(theta+delta))
-lambda.s*(1+((theta+delta)-delta.t)*alpha)
delta.l.r~dt(delta.l, 0.0164,12)
msg<-5*tau.D
dbar~dnorm(delta.l.r,msg)
```

```
pg<-tau.D/2
ssq<-(n-1)*s.y*s.y
ssq~ dgamma(df,pg)
}
```

برای $\bar{d}_\lambda = ۲۱۵$ و $s_Y = ۱۳$، این کد WinBUGS میانگین پسین λ را nm ۵۰۰۰۰۸۳۷ و با انحراف استاندارد پسین nm ۳۴ به‌دست می‌آورد.

بازه اطمینان٪۹۵ نیز (nm ۵۰۰۰۰۹۰۸ nm ۵۰۰۰۰۷۶۸) است. این نتایج هم‌چنین در استاندارد GUM نیز برابر و یکسان هستند.

۱۱-۳-۶ در این راه‌حل، مدل اندازه‌گیری در مورد λ که معادله (۳۱) است، هرگز مورد استفاده قرار نمی‌گیرد، بنابراین باید از کار دشوار و غیرضروری تعیین این‌که چگونه توزیع‌های پارامترهای مختلف باهم مرتبط هستند اجتناب کرد. همانند مثال ۱ با دو پارامتر، رویکرد داده شده در این‌جا منجر به یک توزیع پسین توأم برای همه هشت پارامتر می‌شود.

۱۱-۳-۷ برای این مثال یک راه‌حل تقریبی بر اساس تقریب سری‌های تیلور را بررسی کنید. در راه‌حل استاندارد GUM، معادله (۳۱) به‌صورت تقریبی به‌شکل زیر است:

$$\delta_\lambda = \lambda - \lambda_s \left[1 - \left[\delta_\alpha\left(\overline{\theta} + \Delta\right) + \alpha_s \delta_\theta\right]\right].$$

یک پارامتر $\eta = \lambda - \delta_\lambda$ را تعریف کنید. با استفاده از تقریب سری‌های تیلور، چگالی احتمال η که می‌تواند به‌وسیله توزیع گاوسی برابر $\eta \sim N\left(50000623 \text{ nm}, 911 \text{ nm}^2\right)$ باشد، تقریب می‌شود. برای ساده شدن، هم‌چنین σ_{D_λ} را به S_{d_λ} تقریب کنید. بنابراین، معادله آماری به شکل زیر به‌دست می‌آید:

$$\overline{D}_\lambda \mid \delta_{\lambda_r} \square N\left[\delta_{\lambda_r}, \frac{(13)^2}{5}\right],$$

$$\delta_{\lambda_r} \mid \delta_\lambda \square N(\lambda - \eta, (7{,}8)^2),$$

$$\lambda \square N(0,\text{c})$$

$$\eta \square N(500000623,\ 911{,}47)$$

برای این مدل، چگالی پسین λ می‌تواند از راه تحلیلی به‌دست آید [۳۳]. ما به‌صورت زیر به‌دست می‌آوریم:

$$\lambda \square N\left[\bar{d} + 50000623, \frac{(13)^2}{5} + (7{,}8)^2 + 911{,}47\right]$$

از آنجایی‌که $\bar{d}_\lambda = ۲۱۵$ nm است ما میانگین پسین λ را nm ۵۰۰۰۰۸۳۸ و انحراف استاندارد پسین آن‌را ۳۱nm، به‌دست می‌آوریم که دوباره به نتایج موجود در استاندارد GUM نزدیک هستیم.

۱۱-۴ رویکرد اتکایی

۱۱-۴-۱ از این مثال برای نشان دادن رویکرد استنباط اتکایی در کاربردهای پیچیده‌تر استفاده می‌شود. تابع اندازه‌گیری در فرمول (۳۱) داده شده است. بر اساس اطلاعات ارائه شده در استاندارد GUM مفروضات زیر را داریم:

الف- مقدار برآورد شده λ_s (مثال، مقدار داده شده در گواهی کالیبراسیون)، با l_s مشخص شده است، و برابر با ۵۰۰۰۰۶۲۳ nm است. عدم‌قطعیت استاندارد مرتبط با برآورد ۲۵ nm با ۱۸درجه آزادی است. با یک فرضیه عادی، کمیت اتکایی (FQ) برای λ_s برابر است با:

$$\bar{\lambda}_s = 50000623 - 25T_{18}. \qquad (۳۷)$$

عبارت (۳۷) از توزیع (۲۲) با $\bar{y}$ = ۵۰۰۰۰۶۲۳ nm، $u(\bar{y})$ = ۲۵ nm و ۱۸درجه آزادی مرتبط با $u(\bar{y})$ به‌دست می‌آید.

ب- هر مقدار اندازه‌گیری شده تکراری به‌عنوان یک امتداد از توزیع نرمال با میانگین δ_λ و انحراف استاندارد σ_{δ_λ} در نظر گرفته می‌شود. میانگین مشاهده شده پنج مقدار اندازه‌گیری شده برابر $\bar{d}_\lambda = 215$ nm است. از یک آزمایش جداگانه، σ_{δ_λ} nm ۱۳ با ۲۴ درجه آزادی برآورد می‌شود. بنابراین $u(\bar{d}_\lambda) = 13/\sqrt{5}$ است در نتیجه یک FQ برای δ_λ برابر است با:

$$\bar{\delta}_\lambda = 215 - 13\, T_{24}/\sqrt{5}. \qquad (۳۸)$$

هم‌چنین، بر اساس گواهینامه کالیبراسیون برای ابزار مقایسه، یک برآورد از δ_{c_r}، صفر بوده و عدم‌قطعیت استاندارد آن ۳٫۹ nm (درجه آزادی۵) است و یک برآورد از $\delta_{c_{nr}}$ صفر با عدم‌قطعیت استاندارد ۶٫۷ nm (درجه آزادی ۸) است. علاوه بر این، خطای مقایسه می‌تواند مستقل از خطای تکراری فرض شود، در نتیجه:

$$\tilde{\delta}_{c_r} = 3{,}9T_5, \qquad (۳۹)$$

و

$$\tilde{\delta}_{c_{nr}} = 6{,}7T_8. \qquad (۴۰)$$

استقلال متقابل بین متغیرهای تصادفی نتیجه فرضیه استاندارد GUM درباره فرآیند اندازه‌گیری است.

پ- اجازه دهید $\bar{\theta}$ را به‌عنوان انحراف از میانگین دمای یک آزمایش از مقدار ۲۰ °C در نظر بگیریم. یک برآورد از $\bar{\theta}$ ‎-۰٫۱ °C با انحراف استاندارد ۰٫۲ °C است. چون استاندارد GUM هیچ‌گونه اطلاعات اضافی که این انحراف استاندارد را بررسی کند ارائه نمی‌دهد، درجه‌های آزادی نامتناهی برای آن فرض می‌شود و بنابراین $\bar{\theta}$ یک رسم از توزیع گاوسی است. بنابراین:

$$\tilde{\bar{\theta}} = -0{,}1 - 0{,}2Z, \qquad (۴۱)$$

که در آن:

Z یک امتداد یا رسم از متغیرتصادفی گاوسی استاندارد است که از تمام متغیرهای تصـادفی دیگـر مسـتقل هستند.

ت- یک FQ برایΔ یک تابع چگالی احتمال دارد که به‌شکل زیر است:

$$g(x) = \frac{2}{\pi\sqrt{1-4x^2}}, \quad -0,5\ °C< x <0,5\ °C. \tag{۴۲}$$

برای ایجاد امتدادهای تصادفی از توزیع معکوس سینوس فرمول (۴۲) مشاهده می‌کنیـد کـه اگـر، U_1 یـک متغیر تصادفی$(0,1)$ یکنواخت است $-cos\ (\pi U_1)/2$ توزیع معکوس سیونس موردنیاز است. بنـابراین، یـک FQ برای Δ ممکن است به‌شکل زیر باشد:

$$\tilde{\Delta}= -cos\ (\pi U_1)/2. \tag{۴۳}$$

ث- یک FQ برای δ_α

$$\tilde{\delta}_\alpha = U_2, \tag{۴۴}$$

است که U_2 یک متغیر تصادفی یکنواخت در بازه $\pm 1 \times 10^{-6}$ °C^{-1} است.

ج- یک FQ برای δ_θ برابر است با:

$$\tilde{\delta}_\theta= U_3, \tag{۴۵}$$

است که U_3 یک متغیر تصادفی یکنواخت در بازه $\pm 0,05$ °C است.

چ- یک FQ برای α_s برابر است با:

$$\tilde{\alpha}_s= 11,5 \times 10^{-6} + U_4, \tag{۴۶}$$

که U_4 یک متغیر تصادفی یکنواخت در بازه $\pm 2 \times 10^{-6}$ °C^{-1} است، می‌باشد.

۱۱-۴-۲ با جایگذاری کمیت‌های اتکایی در عبارت‌های (۳۷) تا (۴۶) در عبارت (۳۱)، یک کمیت اتکایی برای λ به‌دست می‌آید. یک تقریب در توزیع برای $\tilde{\lambda}$ با استفاده از ۵۰۰۰۰۰ آزمایش مونت کارلو ارائه می‌شود. میانگین و انحراف استاندارد این توزیع تقریبی به ترتیب ۵۰۰۰۰۸۳۸ nm و ۳۵ nm هستند. بازه اتکایی ۹۵٪ برای λ به‌وسیله کمیت‌های ۰٫۰۲۵ و ۰٫۹۷۵ از این توزیع داده می‌شود. به‌عنوان مثال، (۵۰۰۰۰۹۰۷ nm ۵۰۰۰۰۷۶۸ nm) کد R برای به‌دست آوردن مقدار ۵۰۰۰۰۰ مقدار از $\tilde{\lambda}$ به‌شکل زیر است.

```
nrun = 500000
lambda.s = 50000623 - 25 * rt(nrun, 18)
delta.lambda = 215 - 13/sqrt(5) * rt(nrun, 24)
delta.cr = 3.9*rt(nrun, 5)
delta.cnr = 6.7*rt(nrun, 8)
theta = rnorm(nrun, -0.1, 0.2)
delta = (-cos(pi*runif(nrun))/2)
delta.alpha=runif(nrun,-10^(-6), 10^(-6))
delta.theta=runif(nrun, -0.05, 0.05)
alpha.s=runif(nrun, (11.5-2)*10^(-6), (11.5+2)*10^(-6))
lambda=(lambda.s*(1 + alpha.s*(theta + delta - delta.theta))+delta.lambda
+delta.cr + delta.cnr)/(1 + (alpha.s + delta.aplha)*(theta+delta))
```

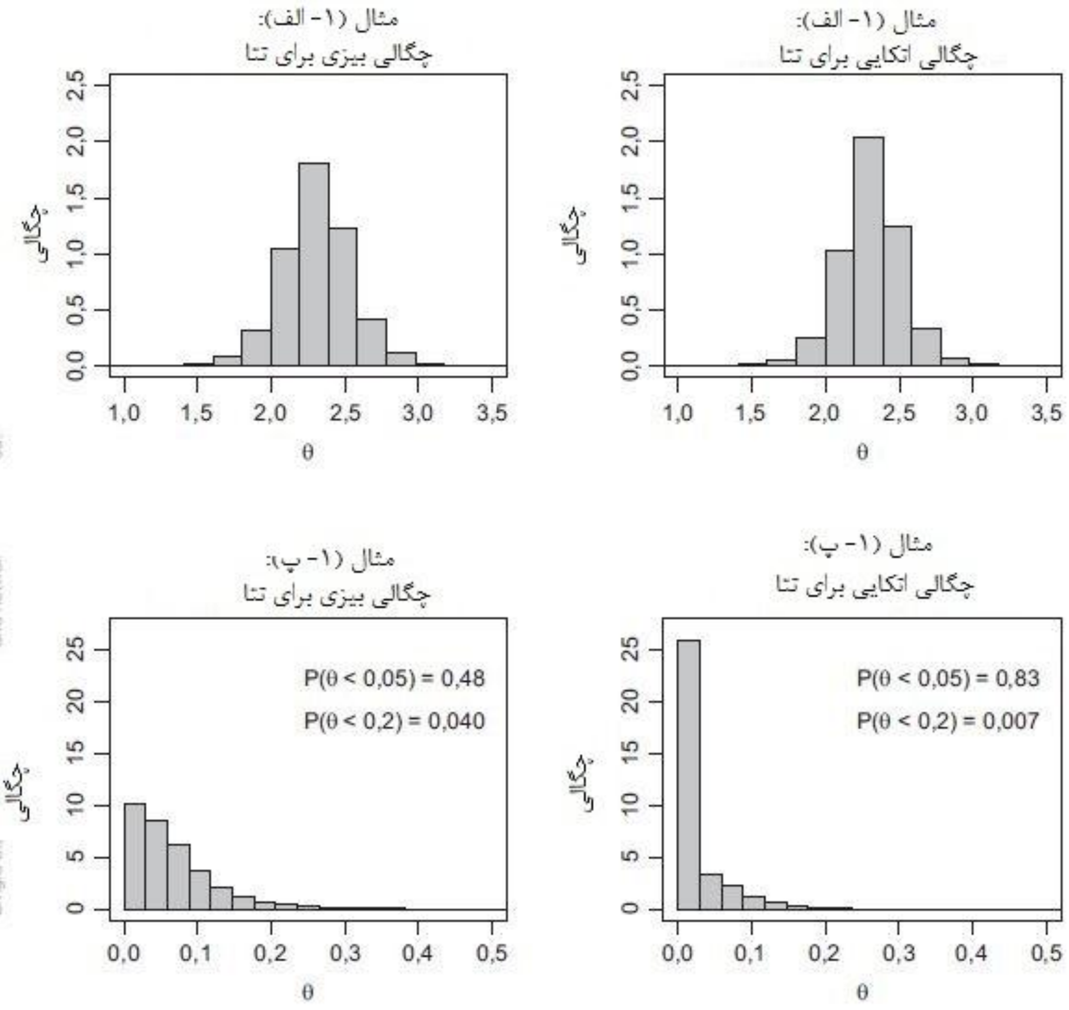
۱۲ بحث

۱-۱۲ مقایسه ارزشیابی‌های عدم‌قطعیت با استفاده از سه رویکرد آماری

۱-۱-۱۲ جدول ۴ نتایج مثال ۱ را به‌طور خلاصه نشان می‌دهد. خودگردان فراوانی‌گرا، راه‌حل‌های بیزی و اتکایی برای مثال (۱- الف) و (۱- ب) خیلی مشابه به هم هستند. راه‌حل‌های خودگردان و استاندارد GUM بازه‌های کوتاه‌تری در هر دو مثال (۱- الف) و (۱- ب) به‌وجود می‌آورند. تفاوت‌های زیاد و قابل توجه در راه‌حل مثال (۱- پ) دیده می‌شود. در این‌جا راه‌حل بیزی بر اساس چگالی پیشین یکنواخت، یک بازه به‌وجود می‌آورد که خیلی بزرگ‌تر از از بازه‌های به‌وجود آمده از سایر روش‌ها است؛ فقط بازه محافظه‌کارانه آیزن‌هارت از آن بزرگ‌تر است.

جدول ۴- بازه‌های بسط یافته عدم‌قطعیت برای سه رویکرد آماری در مثال ۱

	GUM	Eisenhart	Bootstrap	Bayes	Fiducial
مثال (۱- الف)	(۱٫۸۹ ۲٫۷۳)	(۱٫۸۹ ۲٫۷۳)	(۱٫۸۳ ۲٫۶۶)	(۱٫۸۱ ۲٫۸۲)	(۱٫۸۶ ۲٫۷۶)
مثال (۱- ب)	(۱٫۹۰ ۲٫۷۲)	(۱٫۷۸ ۲٫۸۴)	(۱٫۸۶ ۲٫۶۴)	(۱٫۸۳ ۲٫۷۹)	(۱٫۸۷ ۲٫۵۷)
مثال (۱- پ)	(۰٫۰۰ ۰٫۱۲)	(۰٫۰۰ ۰٫۲۰)	(۰٫۰۰ ۰٫۱۱)	(۰٫۰۰ ۰٫۱۹)	(۰٫۰۰ ۰٫۱۴)



شکل۲- مقایسه چگالی های اتکایی و بیزی تقریبی برای مثال (ا- الف) و (۱- پ)

۱۲-۱-۲ چون رویکردهای اتکایی و بیزی، توزیع‌های احتمالی برای اندازه‌ده به‌وجود می‌آورند (θ)؛ نتایج آن‌ها برای مثال (۱- الف) و (۱- پ) به‌طور مفصل‌تر در شکل ۲ مقایسه شده است، که علاوه بر مقایسه بازه‌های عدم‌قطعیت گسترده در جدول ۴ است. نتایج برای مثال (۱- ب) نشان داده نشده‌اند زیرا آن‌ها از نظر شهودی از نتایج مثال (۱- الف) قابل تشخیص نیستند. از نمودارهای شکل۲ واضح است که توزیع احتمال پسین بیزی برای θ و توزیع اتکایی برای θ زمانی که سیگنال به‌طور محسوسی فراتر از زمینه می‌باشد، شبیه به هم هستند. به هر حال، دو توزیع به‌خاطر روش‌های متفاوت آن‌ها در ترکیب محدودیت‌های فیزیکی در مسأله، ویژگی‌های خیلی متفاوتی از خود نشان می‌دهند.

۱۲-۱-۳ در محیط فراوانی‌گرا، اندازه‌ده θ و کمیت ورودی $\mu_1,...,\mu_p$ در مدل اندازه‌گیری (۱) فرض می‌شود که مقادیر ثابت نامعلوم هستند. اگر اندازه‌ده یک ثابت فیزیکی را ارائه کند به‌نظر می‌رسد این رویکرد منطقی باشد به‌طوری که این عمل در مطالعات قبلی یک توزیع پیشین مناسب (اطلاعاتی) و یا یک معادله ساختاری را ارائه نمی‌دهد. آمارشناسان اعتقاد ندارند که تمام پارامترها به‌عنوان متغیرهای تصادفی مدل‌سازی شوند، اگر چه معمولاً با عدم‌قطعیت‌های به‌دست آمده به‌وسیله روش‌های نوع B ارزشیابی و با اختصاص یک توزیع احتمال و انتگرال‌گیری روی آن توزیع سر و کار داشتند. در این مورد، این کار شبیه رویکرد بیزی است، که در آن تمام پارامترها به‌وسیله توزیع‌های احتمال مشخص می‌شوند، اما مفروضات توزیعی کمتری نیاز دارند.

۱۲-۱-۴ خودگردان، یک روش آماری دارای ساختار مناسب است که می‌تواند بازه‌های اطمینان تقریبی غیردقیق و پیچیده را با شبیه‌سازی‌های رایانه‌ای جایگزین کند. طرح‌ها و برنامه‌های خودگردان مختلفی وجود دارد که جهت به‌دست آمدن بازه‌های اطمینان تحت شرایط مختلف توسعه یافته‌اند. بازه خودگردان از *t*- پارامتری، که در این استاندارد معرفی شد، یک انتخاب طبیعی جهت بهبود و گسترش بازه *t*- استیودنت در استاندارد GUM است. مزیت خودگردان کردن، ساده بودن آن است. کاربرد آن برای به‌دست آوردن بازه‌های اطمینان همان‌طوری که در مثال‌ها نشان داده شد سرراست و ساده است.

۱۲-۱-۵ با مثال‌ها نشان داده شد که ارزشیابی عدم‌قطعیت بیزی با استفاده از یک مدل آماری از لحاظ مفهومی ساده است و می‌تواند در مسائل اندازه‌گیری پیچیده بدون هیچ تغییری در روش اصلی به‌کار گرفته شود. اثرات نظام‌مند، نمی‌توانند از مقادیر اندازه‌گیری شده برآورد شوند (که هیچ تابعی از مشاهدات وجود ندارد که مقادیر موردانتظار آن با اثر نظام‌مند برابر باشد) و اطلاعات آن‌ها که برای اجرای ارزشیابی عدم‌قطعیت نوع B مورد استفاده قرار می‌گیرد، می‌توانند به سادگی در مدل بیزی وارد شوند. محاسبه توزیع‌های پسین می‌توانند با استفاده از روش‌های زنجیره مارکوف و مونت کارلو (MCMC) انجام شوند که اغلب نرم‌افزار آن موجود است. همان‌طوری‌که دیدیم هیچ‌گونه بحث مربوط به مجانب‌ها جهت اثبات حالات احتمال وجود ندارد چون نمونه‌های کوچک و بزرگ یک توجیه احتمالی مشترک را نشان می‌دهند.

۱۲-۱-۶ برخی از معایب روش‌های بیزی در اینجا شرح داده شده است. از همه مهم‌تر توزیع‌هـای پیشـین هستند که برای تمام پارامترها در مـدل انـدازه‌گیـری از جملـه انـدازه‌ده، مشـخص مـی‌شـوند. اگـر چـه در اندازه‌شناسی توزیع‌های پیشین اطلاعاتی اغلب به شکل ارزشیابی‌های عدم‌قطعیت نوع B وجود دارند، معمولاً یک یا دو پارامتر موردنیاز است تا به‌صورت توزیع‌های پیشین نامعلوم اختصاص داده شوند (غیراطلاعاتی) که این کار به علت کمبود دانش و اطلاعات پیشین است. چنین توزیع‌های منحصربه‌فرد نیستند، و همان‌طوری که در مثال (۱- پ) نشان داده شد، می‌توانند بر روی نتایج اثر بگذارند. بنابراین توصیه می‌شود برای تحلیل حساسیت به‌منظور اندازه‌گیری مقدار و شـدت چنـین اثراتـی بررسـی شـوند. اثـرات بـزرگ از ویژگـی‌هـای توزیع‌های پیشین غیراطلاعاتی روی می‌دهند و نیاز به مطالعه بیشتر نظام‌های اندازه‌گیری دارنـد. در حالـت کلی وجود چنین اثراتی به این معنی است که اطلاعات ناکـافی در مـورد انـدازه‌ده داده‌هـا وجـود دارد و در نتیجه توزیع پیشین تأثیر قابل ملاحظه‌ای بـر روی نتـایج دارد. در بعضـی از مـوارد، ایـن مشـکل مـی‌توانـد به‌وسیله افزایش تعداد مقادیر اندازه‌گیری شده تکراری یا تغییر روش که در آن داده‌های جمع‌آوری شده‌اند، حل شود به‌عنوان مثال به‌وسیله بهبود راه‌حل. در وضعیت‌های دیگر، ممکن است که یک مدل ریاضی مـورد استفاده قرار گیرد که پارامترهای زیادی دارد و هیچ‌گونه اطلاعات پیشین واقعی در مورد آن پارامترها وجود ندارد، بنابراین این مدل بهتر است به‌صورت ساده درآید.

۱۲-۱-۷ زمانی‌که اطلاعات پیشین معتبر و قابل تـوجهی در مـورد انـدازه‌ده‌هـا وجـود داشـته باشـد، بایـد به‌صورت ساده بیان شده، و به‌صورت مؤثری باید ازطریق قضیه بیزی به‌روزرسـانی شـود. بعـلاوه، حساسـیت نسبت به شکل پیشین، نه فقط برای اندازه‌ده بلکه هم‌چنین برای انحـراف اسـتاندارد از تـابع درسـت‌نمـایی، می‌تواند به‌خوبی نشان داده شود که مشکلی در نظام اندازه‌گیری وجود دارد. این موارد می‌توانند بعـداً مـورد مطالعه قرار گرفته و اصلاح شوند.

۱۲-۱-۸ استنباط اتکایی چارچوبی برای ارتباط یک توزیع با یک پارامتر مورد بررسی را ارائه می‌دهد. نتایج تحقیقات اخیر نشان داد [۲۶] که استنباط اتکایی یک روش آماری معتبر با ویژگیهای عملیاتی خوب اسـت. مثال‌های مورد استفاده نشان دادند که رویکرد اتکایی می‌تواند به سادگی و به‌طور طبیعی اطلاعات مربوط به عدم‌قطعیت را در مدل اندازه‌گیری بگنجاند، و برآوردی از اندازه‌ده و عدم‌قطعیـت اسـتاندارد مـرتبط بـا آن‌را به‌وسیله افزایش توزیع‌های آماری و اجزای آن به‌دست آورد. هیچ نیازی برای نشر عدم‌قطعیت بر اساس بسط سری‌های تیلور یا تقریب ویلش ساترثویت در رویکرد اتکایی وجود ندارد.

۱۲-۱-۹ در این‌جا یک موضوع به‌نام منحصربه‌فرد نبودن در توزیع اتکایی به‌خاطر انتخاب یک شکل خاص از معادله ساختاری وجود دارد. به هرحال، مهم است توجه شود که در بسیاری از کاربردهای علمی، فرآیند فیزیکی به‌وسیله آن داده‌ها به‌وجود می‌آید، معلوم است. در این مورد، معادله ساختاری بهتر است طوری انتخاب شود که این فرآیند را منعکس کند، بنابراین مسأله منحصربه‌فرد نبودن از بین می‌رود. در اندازه‌شناسی، در جائی که یک اندازه‌ده نامعلوم با استفاده از تعدادی فرآیند معلوم اندازه‌گیری می‌شود، بدیهی است، که خطاهای تصادفی در جاهای مشخصی بر برخی از اندازه‌گیری‌ها تأثیر می‌گذارد. مقادیر اندازه‌گیری شده حاصل به‌شکل یک مدل اندازه‌گیری که کمیت‌های اندازه‌گیری شده را با خطاها به‌شکل کمیت‌های اثرگذار ترکیب می‌کند، نشان می‌دهد. این مدل می‌تواند به‌صورت معادله ساختاری باشد.

۱۲-۲ ارتباط بین‌روش‌های پیشنهادی در پیوست ۱ استاندارد GUM (GUMS1) و سه رویکرد آماری

۱۲-۲-۱ GUMS1 [۳] نتایج تصادفی از یک توزیع احتمال را برای یک کمیت خروجی Y در یک مدل اندازه‌گیری که اطلاعات آن کمیت بر اساس اطلاعات کمیت‌های ورودی را توصیف می‌کند، ارائه می‌دهد که به‌وسیله PDFs اختصاص داده شده برای آن‌ها توضیح داده شد. (صفحه vii منبع [۳]). در همان صفحه، GUMS1 می‌گوید که PDF برای یک کمیت به‌عنوان یک چگالی تکراری در نظر گرفته نمی‌شود. در خاتمه در صفحه ۷، می‌گوید که کمیت خروجی Y یک اندازه‌ده است. بنابراین، نتایج تحلیل GUMS1 مانند میانگین و انحراف استاندارد نظریات مونت‌کارلو، برآورد ویژگی‌های چگالی احتمال برای اندازه‌ده است. از این‌رو، یک مقایسه بین نتایج GUMS1 و روش اتکایی و همین‌طور روش مرسوم بیزی نیز امکان‌پذیر است. بازه‌های عدم‌قطعیت GUMS1 احتمالاً به‌خاطر پوشش فراوانی‌گرا مورد مطالعه قرار می‌گیرند اما توصیه می‌شود به‌عنوان بازه‌های اطمینان فراوانی‌گرای عادی تفسیر نشوند.

۱۲-۲-۲ همان‌طوری که در زیربند ۹-۱-۱ و ۹-۱-۲ گفته شد روش‌های مرسوم بیزی مبتنی بر یک مدل آماری است که با اطلاعات پیشین در مورد اندازه‌ده مطابقت دارد. این گفته در مورد GUMS1 صحیح نیست زیرا این روش مبتنی بر یک مدل آماری است، که در آن اندازه‌ده یک کمیت خروجی است، و بنابراین توزیع احتمال آن به‌طور کامل به‌وسیله چگالی‌های احتمال کمیت‌های ورودی تعیین می‌شود. (این قسمت در صفحات ۲ و ۸ منبع [۳] نیز گفته شد) بنابراین هر گونه مقایسه مستقیم بین نتایج حاصل از روش‌های بیزی مرسوم و روش‌های GUMS1 محدود به حالتی است که هیچ‌گونه اطلاعات پیشین از اندازه، موجود نیست.

۱۲-۲-۳ منبع [۳۴] یک چنین مقایسه‌ای را برای مسائل اندازه‌گیری قابل کاربرد گسترده و خاص، انجام می‌دهد. در این استاندارد، اندازه‌ده μ یک تابع از α و β است که از مدل اندازه‌گیری $\mu = f(\alpha,\beta)$ می‌باشد. پارامتر α می‌تواند از داده‌ها برآورد شود زیرا X همان میانگین متغیر تصادفی در توزیع گاوسی است که در آن یک سری مقادیر مشاهده شده وجود دارد. هیچ داده‌ای برای برآورد β وجود ندارد، اما یک توزیع اعتقادی داده می‌شود. تحلیل GUMS1 (به زیربند ۶-۴-۹-۲ در منبع [۳] مراجعه شود) یک توزیع t- استیودنت مقیاسی و انتقال یافته به α نسبت می‌دهد و بنابراین توزیع‌ها را برای α و β با استفاده از تابع f اختصاص می‌دهد. در منبع [۳۴]، نشان داده می‌شود که این تحلیل معادل محاسبه بیزی از چگالی احتمال تابع $f(\alpha,\beta)$ است، که در آن دو پارامتر مستقل هستند، تابع درست‌نمایی برای X ، همان توزیع گاوسی با میانگین α است، و یک توزیع پیشین یکنواخت ناسره برای α استفاده می‌شود و چگالی β برای یک توزیع اعتقادی داده می‌شود. توجه داشته باشید که یک چگالی پیشین برای μ در اینجا مورد استفاده قرار نمی‌گیرد.

۱۲-۲-۴ اکنون فرض کنید که یک تابع g وجود دارد که $\alpha = g(\mu,\beta)$ است. تحلیل مرسوم روش بیزی از تابع درست‌نمایی گاوسی برای X با میانگین $g(\mu,\beta)$ و توزیع‌های پیشین برای μ و β استفاده می‌کند. در نبود هیچ‌گونه اطلاعات اضافی در مورد اندازه‌ده، توزیع یکنواخت ناسره می‌تواند برای μ استفاده شود، اما انتخاب‌های دیگری نیز وجود دارد. توزیع اعتقادی β یک انتخاب طبیعی برای توزیع پیشین است. معمولاً، μ و β به‌صورت متغیرهای تصادفی مستقل در نظر گرفته می‌شوند. توجه کنید که برای این مدل یک چگالی پیشین برای α استفاده نمی‌شود.

۱۲-۲-۵ تحلیل مرسوم بیزی و GUMS1 از پارامتری کردن متفاوتی برای یک مدل مشابه استفاده می‌کنند. نادیده گرفتن اندازه‌ده‌ها تحت هر یک از پارامتری کردن‌ها به‌صورت متفاوتی بیان می‌شود. مدلی که به‌وسیله GUMS1 استفاده می‌شود این کار را مستقیماً به‌وسیله یک چگالی برای μ انجام نمی‌دهد، و به‌جای آن از یک تابع پیشین ناسره غیراطلاعاتی برای α استفاده می‌کند. همان‌طوری که در منبع [۳۴] گفته شد، این یک فرضیه متفاوت است. تحلیل مرسوم بیزی در حالت کلی از یک تابع پیشین غیراطلاعاتی برای خود اندازه‌ده μ استفاده می‌کند. منبع [۳۴] نشان می‌دهد که تحلیل توزیع‌های احتمال یکسانی برای اندازه‌ده زمانی‌که تابع پیشین یکنواخت ناسره برای μ در تحلیل بیزی استفاده می‌شود، به‌وجود می‌آورد که تابع f نیز خطی است. برای توابع غیرخطی توزیع‌های احتمال برای μ تحت ۲ پارامتری کردن که مشابه هم نیستند به‌وجود می‌آیند. مهم است توجه کنیم که اگر توزیع پیشین غیراطلاعاتی برای α به یک توزیع پیشین برای μ منتقل شود بنابراین تحلیل بیزی مرسوم متناظر همان نتایج مشابه GUMS1 را برای هر تابع به‌وجود می‌آورد.

۱۲-۲-۶ همان‌طوری که قبلاً گفته شد بر اساس مدل اندازه‌گیری،GUMS1 یک PDF را برای اندازه‌ده به‌وسیله توزیع PDFs برای کمیت‌های ورودی به‌دست می‌آورد. PDF حاصل اطلاعات اندازه‌ده داده شده بر اساس داده‌ها و مفروضات به‌وجود آمده از تخصیص PDF مشترک برای ورودی را توصیف می‌کند. در بسیاری از مدل‌های استاندارد مقادیر اندازه‌گیری شده بر اساس نتایج به‌دست آمده از توزیع‌های نرمال تک‌متغیره، بازه‌های عدم‌قطعیت با استفاده از روش‌های GUMS1 و اتکایی به‌دست آمدند، که اگر یکسان هم نباشند، بسیار مشابه هم بودند. مدل اندازه‌گیری در مثال (۱- الف) را به‌خاطر بیاورید، به‌عنوان مثال:

$$\theta = \gamma - \beta$$

با $Y_{\mathrm{i}} \sim N\left(\gamma, \sigma_{\gamma}^{2}\right), i = 1, \ldots, 5$ و $B_j \sim N\left(\beta, \sigma_{B}^{2}\right), \mathrm{j} = 1, \ldots, 5$. بر اساس راهنمای GUMS1 یک توزیع مقیاسی و تغییر یافته توزیع t به‌عنوان یکPDF برای γ و β اختصاص می‌دهد به‌ویژه، PDF برای γ همان توزیع مشابه را به‌عنوان متغیر تصادفی دارد.

$$\bar{y} - \frac{s_y}{\sqrt{5}} T_4^{(1)},$$

که $T_4^{(1)}$ یک متغیر تصادفیt- استیودنت با ۴ درجه آزادی است و PDF برای β همان توزیع مشابه را به‌عنوان متغیر تصادفی دارد.

$$\bar{b} - \frac{s_b}{\sqrt{5}} T_4^{(2)},$$

که $T_4^{(2)}$ یک متغیر تصادفی t- استیودنت با ۴ درجه آزادی است که مستقل از $\mathrm{T}_4^{(1)}$ است. در نتیجه، PDF برای اندازه‌دهθ می‌تواند از از توزیع

$$\bar{y} - \bar{b} - \frac{s_y}{\sqrt{5}} T_4^{(1)} + \frac{s_b}{\sqrt{5}} T_4^{(2)}$$

به‌دست آید.

کد R برای به‌دست آوردن ۵۰۰۰۰۰ مقدار از توزیع فوق به‌شکل زیر است:

```
nrun = 500000
T1 = rt(nrun, 4)
T2 = rt(nrun, 4)
theta = 3.537 - 1.228 - 0.342/sqrt(5)*T1 + 0.131/sqrt(5)*T2
```

یک بازه عدم‌قطعیت٪۹۵ در ۰/۰۲۵ و ۰/۹۷۵ از تقریب PDF برابر است با:

```
quantile(theta, c(0.025, 0.975))
## 2.5% 97.5%
## 1.853703 2.763999
```

که اساساً با بازه اتکایی برای این مثال که قبلاً داده شده بود یکسان است. به‌طور مشابه، GUMS1 و رویکردهای اتکایی، بازه‌های عدم‌قطعیت مشابهی را در مثال‌های (۱- ب) و (۱- پ) به‌وجود می‌آورند.

۱۲-۲-۷ در بسیاری از موارد دیگر، روش‌هایGUMS1 و اتکایی نتایج متفاوتی را به‌وجود می‌آورند. یک مورد"حداکثر" می‌تواند در مسأله گفته شده درمنبع [۳۵] پیدا شود. در مثال ارائه شده در این منبع [۳۵] اندازه‌ده، مقدار و حجم یک کمیت مقداری پیچیده دارند.

$$\Gamma = \Gamma_1 + iT_2.$$

که اندزه‌ده آن، برابر است با:

$$|\Gamma| = \sqrt{\Gamma_1^2 + \Gamma_2^2}.$$

با فرض این‌که $X_1 \sim N(\Gamma_1, \sigma^2)$ و $X_2 \sim N(\Gamma_2, \sigma^2)$ با σ معلوم،GUMS1 که $N(x_1, \sigma^2)$ PDF را برای Γ_1 و $N(x_2, \sigma^2)$ را برای Γ_2 تعیین می‌کند. در نتیجه،PDF برای $|\Gamma|$ براساس GUMS1 یک توزیع برای متغیر تصادفی

$$\sqrt{(x_1 - \sigma Z_1)^2 + (x_2 - \sigma Z_2)^2}, \qquad (۴۳)$$

است. که Z_2 و Z_1 متغیرهای تصادفی نرمال استاندارد مستقل هستند. منبع [۳۵] نشان داد که بازه‌های GUMS1 برای $|\Gamma|$ زمانی که $|\Gamma|$ در مقایسه با σ کوچک باشد، عملکرد فراوانی‌گرا (احتمالات پوششی ناکافی) نامطلوب دارد. این کار به این علت است که متغیر تصادفی در عبارت (۴۳) مثبت است و بنابراین کران پایین بازه عدم‌قطعیت برای $|\Gamma|$ نیز مثبت خواهد بود، که ممکن است زمانی که $|\Gamma|$ نزدیک به صفر است آن‌را پوشش ندهد.

۱۲-۲-۸ یک راه‌حل اتکایی برای مسأله می‌تواند بر اساس این واقعیت که $\left(X_1^2+X_2^2\right)/\sigma^2$ به‌عنوان توزیع λ^2 غیرمرکزی با ۲ درجه آزادی و غیرمرکزی بودن پارامتر $\lambda = |\Gamma|^2\sigma^2$ باشد. این ویژگی توزیعی می‌تواند برای بسط و توسعه یک معادله ساختاری که آماره قابل مشاهده $\left(X_1^2+X_2^2\right)/\sigma^2$ را به λ مربوط می‌کند مورد استفاده قرار گیرد، که شامل پارامتری از $|\Gamma|$ است. بر اساس این معادله ساختاری، یک بازه اتکایی برای $|\Gamma|$ می‌تواند ساخته شود. منبع [۳۶] نشان داد که این بازه اتکایی پوشش فراوانی‌گرای اسمی را در تمام وضعیت‌ها حفظ می‌کند.

۱۳ خلاصه

۱-۱۳ در این استاندارد، سه رویکرد برای ساخت بازه‌های عدم‌قطعیت که تفسیرهای احتمالی شفافی داشتند مورد بحث قرار گرفت. در مقابل، کارهای زیاد دیگری در این زمینه بر روی تعیین اجزای آماری رویه‌های که در حال حاضر در این حوزه در جامعه اندازه‌شناسی مورد استفاده قرار می‌گیرد، متمرکز شده است. یکی از اهداف در رویکرد مطالعه روش‌ها، برای ارزشیابی عدم‌قطعیت از یک چنین منظری تلاش برای به‌دست آوردن یک درک از روش‌های حاضر و تأکید بر گزینه‌های جدیدی است که ممکن است مفید باشند.

۲-۱۳ همان‌طوری که در منبع [۹] مشاهده شد بازه‌های عدم‌قطعیت که تحت رویکردهای مختلف به‌وجود آمدند، اغلب از لحاظ عددی شبیه هم خواهند بود. حتی در این مورد، هر چند، تفسیر آن‌ها متفاوت است.

۳-۱۳ بازه‌های عدم‌قطعیت فراوانی‌گرا حالت‌های احتمالی را در مورد عملکرد بلندمدت یک رویه خاص برای ساخت و تشکیل بازه‌های عدم‌قطعیت تحت شرایط یکسان به‌وجود می‌آورد. بنابراین، حالت احتمالی مستقیماً در مورد اندازه‌ده، بیان نمی‌شود اما در مورد ارتباط بلندمدت بین یک رویه با این‌که یک بازه و یک اندازه‌ده به‌وجود آمده است، به‌کار می‌رود. ابتدا مقادیر اندازه‌گیری شده به‌دست می‌آیند و یک بازه عدم‌قطعیت فراوانی‌گرا محاسبه می‌شود، در این‌جا هیچ‌چیز تصادفی در مورد نتایج وجود ندارد. اگرچه معلوم نیست که آیا اندازه‌ده در هر بازه خاص به‌دست آمده است یا نه، چنین بازه‌هایی اندازه‌ده را با یک احتمال خاص به‌دست خواهند آورد. برخلاف بازه عدم‌قطعیت متداول فقط بر اساس اطلاعات آماری، بازه عدم‌قطعیت فروانی‌گرا به‌طور عادی محاسبه می‌شود و در نتیجه سطح اطمینان مطلوب به‌طور متوسط با انتگرال‌گیری روی توزیع احتمال هر کمیت که با استفاده از ارزشیابی نوع B به‌دست می‌آیند، محاسبه می‌شود.

۴-۱۳ بازه‌های اتکایی و بیزی از طرف دیگر بر اساس توزیع‌های احتمال قرار دارند که به‌صورت مستقیم اطلاعات اندازه‌ده را توصیف می‌کنند. روش‌های استفاده شده برای به‌دست آوردن این دو نوع بازه متفاوت هستند، اما نتایج آن‌ها در این بُعد از تفسیر مشابه خواهد بود. نتایج بیزی با ترکیب توزیع‌های احتمال برای هر پارامتر پیشین و خاص و تحلیل داده‌ها با یک مدل احتمال به‌دست می‌آیند که تغییر داده‌هایی را که از قضیه بیزی استفاده می‌کنند را توصیف می‌کند. نتیجه‌گیری از توزیع‌های پسین برای هر پارامتر احتمال مقادیر پارامتر داده شده با توجه به اطلاعات پیشین و داده‌ها را منعکس می‌کند. نتایج اتکایی به‌وسیله معکوس کردن یک مدل احتمالی که از داده‌ها به پارامترها تعریف شده است، به توزیعی می‌باشد که از پارامترها به داده‌ها تعریف می‌شود.

۱۳-۵ اگر نتایج عددی همیشه شبیه به‌هم باشند، هر یک از تفسیرها در هر بازه عدم‌قطعیت قابل کاربرد خواهد بود (حداقل به‌طور تقریبی). همان‌طوری که در مثال‌های این استاندارد نشان داده شد، نتایج عددی می‌توانند در بعضی از موارد از همدیگر متفاوت باشند. اگرچه، هر کدام از آنها می‌توانند از لحاظ احتمالی تأیید شوند و سطح مشترکی از اهمیت را به اشتراک بگذارند. (در حالت کلی۹۵٪). تفاوت‌های دیگری نیز ممکن است مشاهده شود. برای مثال اگر یکی از منابع مهم عدم‌قطعیت در یک کاربرد خاص به یک کمیت دارای توزیع مبهم باشد، بازه عدم‌قطعیت به‌دست آمده با استفاده از رویکردهای بیزی یا اتکایی، این عدم‌تقارن را نشان می‌دهند، درحالی‌که یک بازه اطمینان تقریبی که با استفاده از رویه‌های GUMS به‌دست آمده است یک بازه عدم‌قطعیت متقارن به‌وجود خواهند آورد. و (ممکن است در یک طرف طولانی‌تر از مقدار موردنیاز باشد). نتایج رویکرد فراوانی‌گرا مبتنی بر سایر اصول آماری ممکن است در بعضی از موارد با نتایج بیزی یا اتکایی همخوانی و مطابقت داشته باشد، اما روش‌های متفاوت در حالت کلی مطابق و هماهنگ نخواهند بود زیرا هر رویکرد در خاتمه مبتنی بر یک سری مفروضات و معیارهای ریاضی مختلف است.

۱۳-۶ وجود رویکردهای مختلف، برای ارزشیابی عدم‌قطعیت که همیشه هماهنگ و سازگار نیستند ممکن است پیچیده به‌نظر برسند. به هر حال، بهتر است به‌عنوان یک نشانه از فرصت‌های بیشتر در نظر گرفته شود. تنها با همکاری دائمی برای درک ویژگی‌های رویکردهای متفاوت است که روش‌های ارزشیابی عدم‌قطعیت به منظور رفع تمام نیازهای علمی و اقتصادی به‌دست می‌آید: روش‌های که برای پیاده‌سازی عملی هستند، از منابع استفاده بهینه می‌کنند، در بسیاری از انواع اندازه‌گیری قابل کاربرد خواهند بود، اندازه‌گیری‌های که (هم قدیمی و هم جدید) بوده و دارای مفهوم شفافی هستند.

کتاب‌نامه

[1] ISO/IEC Guide 98-3:2008, Uncertainty of measurement — Part 3: Guide to the expression of uncertainty in measurement (GUM:1995)

[2] ISO/IEC 17025:2005, General requirements for the competence of testing and calibration laboratories,

یادآوری- استاندارد ملی ایران شماره ۱۷۰۲۵: سال۱۳۸۶، الزامات عمومی برای احراز صلاحیت آزمایشگاه های آزمون و کالیبراسیون با استفاده از استاندارد بین المللی ISO/IEC 17025 : 2005 تدوین شده است.

[3] ISO/IEC Guide 98-3:2008/Suppl 1:2008, Uncertainty of measurement — Part 3: Guide to the expression of uncertainty in measurement (GUM:1995) — Supplement 1: Propagation of distributions using a Monte Carlo method

[4] ISO/IEC Guide 99:2007, International vocabulary of metrology — Basic and general concepts and associated terms (VIM)

یادآوری- استاندارد ملی ایران شماره ۴۷۲۳: سال۱۳۹۰، واژه‌نامه اندازه‌شناسی- مفاهیم پایه و عمومی و اصطلاحات مربوط با استفاده از استاندارد بین المللی ISO/IEC Guide 99: 2007 تدوین شده است.

[5] GLESER, L.J. Assessing uncertainty in measurement. Statistical Science, 13:277–290, 1998

[6] KACKER, R. and JONES, A. On use of Bayesian statistics to make the guide to the expression of uncertainty in measurement consistent. Metrologia, 40:235–248, 2003

[7] ELSTER, C. W., WÖGER, W. and COX, M.G. Draft GUM Supplement 1 and Bayesian analysis. Metrologia, 44:L31–L32, 2007

[8] WILLINK, R. A procedure for the evaluation of measurement uncertainty based on moments. Metrologia, 42:329–343, 2005

[9] LIRA, I. and WÖGER, W. Comparison between the conventional and Bayesian approaches to evaluate measurement data. Metrologia, 43: S249–S259, 2006

[10] GUTHRIE, W.F., LIU, H.K., RUKHIN, A.L., TOMAN, B., WANG, C.M. and ZHANG, N.F. Three statistical paradigms for the assessment and interpretation of measurement uncertainty. Data Modeling for Metrology and Testing in Measurement Sciences, edit. Pavese, F. and Forbes, A. B. Birkhauser, Boston, 2008

[11] KIRKUP, L. and FRENKEL, B. An Introduction to Uncertainty in Measurement Using the GUM. Cambridge University Press, Cambridge UK, 2006

[12] BAYES, T. An essay toward solving a problem in the doctrine of chances. Philosophical Transactions of the Royal Society of London, 53:370–418, 1764. (facsimile available at http://www.stat.ucla.edu/history/essay.pdf)

[13] FISHER, R.A. Inverse probability. Proc. Comb. Philos. Soc., 26:528–535, 1930

[14] CASELLA, G. and BERGER, R. Statistical Inference. Duxbury, MA, 2 edition, 2002

[15] GUTHRIE, W.F. Should (T1-T2) have larger uncertainty than T1? Proceedings of the 8th International Conference on Temperature: Its Measurements and Control, 2:887–892, 2002. http://www.itl.nist.gov/div898/pubs/author/guthrie/guthrie-2002-01.pdf)

[16] EFRON, B. and TIBSHIRANI, R.J. An Introduction to the Bootstrap. Monographs of Statistics and Applied Probability, volume 57. Chapman and Hall, 1993

[17] EISENHART, C. Expression of the uncertainties of final measurements results. NBS special publication, NIST, Gaithersburg, MD, 1983

[18] R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, 2003. ISBN 3-900051-00-3, http://www.R-project.org

[19] LUNN, D.J, THOMAS, A., BEST, N. and SPIEGELHALTER, D. WinBUGS – a Bayesian modeling framework: concepts, structure, and extensibility. Statistics and Computing, 10:325-337, 2000

[20] BERNARDO, J.M. and SMITH, A.F.M. Bayesian Theory. John Wiley and Sons Ltd,1994

[21] HASTIE, T., TIBSHIRANI, R. and FRIEDMAN, J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer-Verlag, New York, 2001

[22] GELMAN, A., CARLIN, J.B., STERN, H.S., and RUBIN, D.B. Bayesian Data Analysis. Chapman and Hall, 1995

[23] BROWNE, W.J. and DRAPER, D. A comparison of Bayesian and likelihood-based methods for fitting multilevel models. Bayesian Analysis, 1:473–514, 2006

[24] WANG, C.M. and IYER, H.K. Propagation of uncertainties in measurements using generalized inference. Metrologia, 42:145–153, 2005

[25] WANG, C.M. and IYER, H.K. A generalized confidence interval for a measurand in the presence of Type- A and Type-B uncertainties. Measurement, 39:856–863, 2006

[26] HANNIG, J., IYER, H. K., and PATTERSON, P. L. Fiducial generalized confidence intervals. Journal of the American Statistical Association, 101:254–269, 2006

[27] WANG, C.M. and IYER, H.K. Uncertainty analysis for vector measurands using Fiducial inference.

Metrologia, 43:486–494, 2006

[28] FRASER, D.A.S. The Structure of Inference. New York: Krieger, 1968

[29] HANNIG, J. On Fiducial inference – the good, the bad and the ugly. Technical Report 2006/3, Department of Statistics, Colorado State University, Fort Collins, CO, 2006.

URL http://www.stat.colostate.edu/research/2006_3.pdf

[30] IYER, H.K. and PATTERSON, P.L. A recipe for constructing generalized pivot quantities and generalized confidence intervals. Technical Report 2002/10, Department of Statistics, Colorado State University,

Fort Collins, CO, 2002. URL http://www.stat.colostate.edu/research/2002_10.pdf

[31] RUKHIN, A.L. and SEDRANSK, N. Statistics in metrology: international key comparisons and interlaboratory studies. Journal of Data Science, 7:393–412, 2007

[32] EFRON, B. Six questions raised by the bootstrap. In: Exploring the Limits of Bootstrap (R. LePage and L. Billard, editors) pages 99–126. Wiley, NY, 1992

[33] LINDLEY, D. and SMITH, A.F.M. Bayes estimates for the linear model, JRSS B., 34:1-41, 1972

[34] ELSTER, C. and TOMAN, B. Bayesian uncertainty analysis under prior ignorance of the measurand versus analysis using the Supplement 1 to the Guide: a comparison, Metrologia, 46:261-266, 2009

[35] HALL, B.D. Evaluating methods of calculating measurement uncertainty, Metrologia, 45:L5-L8, 2008

[36] WANG, C.M. and IYER, H.K. Fiducial intervals for the magnitude of a complex-valued quantity, Metrologia, 46:1 81-86, 2009