

۲۴۸۷۶۶۶

مبانی استنتاج آماری

معنای خطای تصادفی چیست؟

نوربرت هیرشاور

سون گرونر

الیور موسوف

www.ketab.ir

مترجم:

دکتر مهدی فتح آبادی

عضو هیئت علمی دانشگاه آزاد اسلامی واحد فیروزکوه

سرشناسه	هیرشاوَر، نوربرت، 1972، Norbert Hirschauer
عنوان و پدیدآور	مبانی استنتاج آماری، معنای خطای تصادفی چیست؟/ نویسنده: نوربرت هیرشاوَر - سون گروَنِر - الیور موسوف؛ مترجم: دکتر مهدی فتح آبادی.
مشخصات نشر	تهران، نور علم.
مشخصات ظاهری	۱۸۶ص.
عنوان اصلی	Fundamentals of Statistical Inference. What is the Meaning of Random Error?
شابک	۹۷۸-۶۰۰-۱۶۹-۶۲۴-۴
موضوع	آمار ریاضی.
	احتمالات.
شناسه افزوده	فتح آبادی، مهدی، ۱۳۶۱، مترجم.
شناسه افزوده	غفاری، فرهاد، ۱۳۴۶، ویراستار علمی.
رده بندی کنگره	QA ۲۷۶
رده بندی دیویی	۵۱۹/۵

نشر نور علم و پخش قلم سینا: تهران - انقلاب - خ ۱۲ فروردین - پلاک ۲۸۶ - تلفن: ۶۶۴۰۵۸۸۰

۶۶۹۵۷۱۲۰-۰۹۱۲۳۴۶۲۸۱۱-۶۶۹۵۷۰۲۱

وب سایت <https://www.modiranketab.ir> پیج اینستاگرام @modiranketab

مبانی استنتاج آماری، معنای خطای تصادفی چیست؟

نویسنده: نوربرت هیرشاوَر - سون گروَنِر - الیور موسوف.

مترجم: دکتر مهدی فتح آبادی - عضو هیئت علمی دانشگاه آزاد اسلامی واحد فیروزکوه

ویراستار علمی: دکتر فرهاد غفاری - دانشیار دانشگاه آزاد اسلامی علوم و تحقیقات تهران

ناشر: نور علم

شمارگان: ۱۰۰ جلد

شابک: ۹۷۸-۶۰۰-۱۶۹-۶۲۴-۴

نوبت چاپ: اول ۱۴۰۳

چاپ و صحافی: سورنا

قیمت: ۱۷۰۰۰۰ تومان

در صورت عدم دسترسی به کتابهای این انتشارات، از طریق تماس با
۰۹۱۲۳۳۳۴۲۲۹ کتابها با پست به تمام نقاط ایران ارسال می شود.

در بیشتر محیط‌های علمی، طبقه‌بندی قراردادی نتایج به «معنادار» و «غیرمعنادار» برای تفسیر معتبر داده‌ها ضروری نبوده و اغلب زیان‌بار است. تخمین اندازه اثر و عدم قطعیت پیرامون برآوردهای ما برای استنتاج علمی و قضاوت صحیح بسیار مهم‌تر از هر طبقه‌بندی مشابه دیگری خواهد بود. گرینلند و همکاران (۲۰۱۶: ۳۳۸): آزمون‌های آماری، مقادیر احتمال (p-values)، فواصل اطمینان و قدرت؛ یک رهنمون برای تفسیرهای نادرست

در زندگی مدرن، گزاره‌های علمی مبتنی بر داده‌ها درباره جهان در همه جا حضور داشته و برای تصمیم‌گیری درست در سازمان‌ها و جامعه به‌عنوان یک کل بسیار مهم هستند. به تغییرات جوی یا همه‌گیری کووید ۱۹ در کنار موضوعاتی هم‌چون نحوه عملکرد ماسک صورت، واکسن و یا محدودیت‌های نظارتی درباره جابجایی افراد فکر کنید. مطالعات درست به ما کمک می‌کنند که به انباشت دانش پرداخته و با توجه به عدم قطعیت‌های موجود، معقول‌ترین گزاره‌ها را درباره حقایق (ناشناخته) دنیای واقعی بیان کنیم. در مقابل، مطالعات ضعیف و همچنین تفسیر نادرست نتایج، مانع قضاوت صحیح شده و می‌تواند به نتیجه‌گیری‌های بیش از حد مطمئن منجر شود؛ که این ما را به بیراهه می‌کشانند.

برای استنتاج بامعنی درباره زمینه‌ای وسیع‌تر براساس یک مجموعه داده محدود (مانند یک نمونه تصادفی که از یک جامعه گرفته شده)، آمارها می‌توانند ابزارهایی مفید باشند. با این حال، به نظر می‌رسد شیوه مرسوم آزمون معناداری آماری جای اینکه به محققان کمک کند تا با در نظر گرفتن عدم قطعیت ناشی از خطای تصادفی، درباره سهم دانش یک مطالعه قضاوت درستی داشته باشند، باعث سردرگمی آنها می‌شود. علی‌رغم انتقادات موجود، بکارگیری و تفسیرهای نادرست از آزمون‌های معناداری آماری به طور نگران‌کننده‌ای همچنان یک عمل «عادی» به شمار می‌آید. اگرچه مباحث علمی درباره این موضوعات در چند سال گذشته گسترش یافته است؛ مانند هشدار روش‌شناختی بی‌سابقه توسط انجمن آمار آمریکا در سال ۲۰۱۶ و سخن از کنار گذاشتن معناداری آماری در سال ۲۰۱۹؛ اما به نظر می‌رسد مشکلات همچنان پابرجا هستند.

این کتاب تلاش می‌کند نگرانی عمیق خود را درباره جریان ادامه‌دار ادعاهای معناداری همراه‌کننده و گزاره‌های بیش از حد مطمئن اکتشافات علمی (نتیجه‌گیری دوتایی بله یا خیر) بیان داشته که به اعتبار علم خدشه وارد کرده و خطر انتخاب‌های ناآگاهانه را افزایش داده است. اما این کتاب از همان ابتدا به این اشتباه‌های استنتاجی نمی‌پردازد؛ بلکه قبل از بحث درباره تفسیرهای نادرست که همراه با گزاره‌های معناداری دوگانه هستند، ابتدا مبانی مفهومی استنتاج آماری بیان می‌گردد. به عبارت دیگر، نشان خواهیم داد در بهترین حالت، آنچه که داریم یک برآورد نارایب مبتنی بر داده‌ها از اندازه اثر (به عنوان مثال، شکاف دستمزد جنسیتی یا رابطه

بین تحصیلات و درآمد در یک جامعه) و هم چنین تخمین «عدم قطعیت ناشی از خطای تصادفی برآوردها» می باشد که این برآوردها نیز مبتنی بر داده ها خواهند بود.

ما چند نوآوری در بحث روش شناختی ارائه کرده ایم تا به این پرسش پاسخ دهیم؛ چگونه می توان از خطاهای استنتاجی در آزمون های معناداری مرسوم جلوگیری نمود. هر یک از این نوآوری ها جنبه های مختلفی از مساله را در بر می گیرند. این کتاب با تکیه بر این موضوع، توصیفی منسجم از مبانی مفهومی استنتاج آماری ارائه داده و نشان می دهد چگونه می توان تغییراتی برای بهتر شدن ایجاد کرد. بدین منظور، ما از مقالات زیر استفاده کرده ایم؛

Hirschauer, N. (2022): Unanswered questions in the p-value debate. Significance (June 2022): 42–44 (<https://doi.org/10.1111/1740-9713.01655>).

Hirschauer, N., Grüner, S., Mußhoff, O., Becker, C., Jantsch, A. (2021b): Inference using non-random samples? Stop right there! Significance (October 2021): 20–24 (<https://doi.org/10.1111/1740-9713.01568>).

Hirschauer, N., Grüner, S., Mußhoff, O., Becker, C. (2021a): A Primer on p-Value Thresholds and α -Levels – Two Different Kettles of Fish. German Journal of Agricultural Economics 70: 123–133 (<https://doi.org/10.30430/70.2021.2.123-133>).

Hirschauer, N., Grüner, S., Mußhoff, O., Becker, C., Jantsch, A. (2020b): Can p-values be meaningfully interpreted without random sampling? Statistics Surveys 14(2020): 71–91 (<https://doi.org/10.1214/20-SS129>).

Hirschauer, N., Grüner, S., Mußhoff, O., Becker, C. (2020a): Inference in economic experiments. Economics. The Open-Access, Open-Assessment E-Journal 14(2020-7): 1–14 (<https://doi.org/10.5018/economics-ejournal.ja.2020-7>).

Hirschauer, N., Grüner, S., Mußhoff, O., Becker, C. (2019): Twenty steps towards an adequate inferential interpretation of p-values in econometrics. Journal of Economics and Statistics 239(4): 703–721 (<https://doi.org/10.1515/jbnst/2018-0069>).

Hirschauer, N., Grüner, S., Mußhoff, O., Becker, C. (2018): Pitfalls of significance testing and p-value variability: An econometrics perspective. Statistics Surveys 12(2018): 136–172 (<https://doi.org/10.1214/18-SS122>).

Hirschauer, N., Mußhoff, O., Gruener, S., Frey, U., Theesfeld, I., Wagner, P. (2016): Inferential misconceptions and replication crisis. Journal of Epidemiology, Biostatistics, and Public Health 13(4): e12066-1–e12066-16 (<https://riviste.unimi.it/index.php/ebph/article/view/17632>).

افراد زیادی با پرسش‌های انتقادی و پیشنهادات خود، ما را در نوشتن این کتاب حمایت کردند. مهمتر از همه، آنها به ما گوشزد کردند که دانشجویان و متخصصان آمار را با سوابق آماری مختلف در خاطر داشته باشیم. ما به‌ویژه از همکارانمان در دانشگاه مارتین لوتر هال-ویتنبرگ و موسسه توسعه کشاورزی لایبنیتس در اقتصادهای در حال گذار در هاله که در جلسات کاری غیررسمی ما درباره استنتاج آماری سؤالات مفیدی پرسیدند و ما را تشویق کردند، سپاسگزاریم. به طور خاص، مایلیم از نویسندگان مقالات مبتنی بر استنتاج که سخاوتمندانه به ما اجازه دادند تا از مطالعه آنها به عنوان منابع این کتاب استفاده کنیم، تشکر کنیم: کلودیا بکر، اولریش فری، آنته جانتش، اینسا تیسفلد، و پیتر واگنر.

Halle (Saale), Germany

Norbert Hirschauer

Halle (Saale), Germany

Sven Grüner

Göttingen, Germany

Oliver Mußhoff

www.ketab.ir

فهرست مطالب

فصل یک	۹
مقدمه	۹
فصل دو	۲۰
معنای استنتاج علمی و آماری	۲۰
۱-۲. شروع بحث: خطاها و ارزیابی روایی	۲۰
۲-۲. روایی خارجی	۲۳
۳-۲. روایی داخلی	۲۷
۴-۲. خلاصه فصل؛ استنتاج علمی فراتر از استنتاج آماری است	۲۹
۵-۲. مطالعات پیشنهادی	۳۱
فصل سه	۳۲
مبانی استنتاج آماری: نمونه‌گیری تصادفی ساده	۳۲
۱-۳. شروع بحث: آماره‌های توصیفی یک مجموعه داده معین	۳۲
۲-۳. نمونه‌گیری تصادفی، خطای نمونه‌گیری و توزیع نمونه‌گیری	۳۴
۳-۳. تخمین و عدم قطعیت تخمین در نمونه‌گیری تصادفی ساده	۳۹
۱-۳-۳. برآورد اندازه اثر و خطاهای استاندارد مبتنی بر نمونه	۳۹
۲-۳-۳. یک مثال کاربردی: شکاف دستمزد جنسیتی	۴۴
۳-۳-۳. تغییرپذیری نمونه به نمونه تخمین‌های نقطه‌ای و خطای استاندارد	۴۶
۴-۳. خلاصه فصل: فروض آماری از الزامات تجربی هستند	۵۱
۵-۳. مطالعات پیشنهادی	۵۲
فصل چهار	۵۴
عدم قطعیت تخمین در طرح‌های نمونه‌گیری پیچیده	۵۴
۱-۴. مروری بر طرح‌های مختلف نمونه‌گیری	۵۴
۲-۴. نمونه‌گیری طبقه‌ای	۵۶
۳-۴. نمونه‌گیری خوشه‌ای	۵۸
۴-۴. نمونه‌های ساده آغشته به تورش انتخاب	۶۴

- ۴-۱. غیر تصادفی بودن: چالشی بزرگ در علوم اجتماعی..... ۶۴
- ۴-۲. رویکردهایی برای تصحیح تورش انتخاب در نمونه‌های ساده..... ۶۶
- ۴-۵. جامعه‌های کامل و تصحیح جامعه محدود..... ۷۱
- ۴-۶. خلاصه فصل: استنتاج مستلزم در نظر گرفتن طرح نمونه‌گیری است..... ۷۵
- ۴-۷. مطالعات پیشنهادی..... ۷۶
- فصل پنج..... ۷۸
- انباشت دانش از طریق فراتحلیل و تکرارها..... ۷۸
- ۵-۱. مبانی فراتحلیل..... ۷۸
- ۵-۱-۱. نحوه تقابل با اندازه‌گیری‌های مختلف و تصریح‌های مدل..... ۷۸
- ۵-۱-۲. ترکیب اندازه‌های اثر و خطاهای استاندارد چندین مطالعه..... ۸۳
- ۵-۲. ارزیابی کیفیت تحقیق از طریق تکرارها..... ۸۸
- ۵-۳. خلاصه فصل: بهترین برآوردهای ما به طور متوسط به درستی تخمین می‌زنند..... ۹۵
- ۵-۴. مطالعات پیشنهادی..... ۹۷
- فصل شش..... ۹۸
- مقدار احتمال (P-Value) و آزمون معناداری آماری..... ۹۸
- ۶-۱. مفهوم مقدار احتمال (P-Value)..... ۹۸
- ۶-۲. آزمون معناداری فرضیه تهی..... ۱۰۴
- ۶-۲-۱. دوگانه‌سازی مقدار احتمال و اظهارنظرهای معناداری..... ۱۰۴
- ۶-۲-۲. آیین آماری «آزمون معناداری فرضیه تهی» و تفسیرهای نادرست از مطالعات انفرادی..... ۱۰۷
- ۶-۲-۳. ماندگاری آیین آماری «آزمون معناداری فرضیه تهی» در مطالعات تکراری..... ۱۱۳
- ۶-۲-۴. لغزش‌ها و تورش انتشار در شیوه «آزمون معناداری فرضیه تهی»..... ۱۱۶
- ۶-۲-۵. رویکردهایی با هدف کاهش تورش انتشار..... ۱۲۳
- ۶-۳. ریشه‌های تاریخی چارچوب «آزمون معناداری فرضیه تهی»..... ۱۲۶
- ۶-۳-۱. آزمون معناداری فرضیه تهی: ترکیبی از دو رویکرد آماری سازش‌ناپذیر..... ۱۲۶

۱۲۸.....	۲-۳-۶. رفتار استقرایی (آزمون فرضیه) و اندازه خطای نوع اول (α)
۱۳۹.....	۳-۳-۶. باور استقرایی (آزمون معناداری) و آستانه‌های مقادیر احتمال (p-Value)
۱۴۲.....	۴-۶. خلاصه فصل: از اظهارنظرهای معناداری باید اجتناب کرد.....
۱۴۵.....	۵-۶. مطالعات پیشنهادی.....
۱۴۷.....	فصل هفت.....
۱۴۷.....	استنتاج آماری در آزمایش‌ها.....
۱۴۷.....	۱-۷. حالت‌های استنتاجی در مقایسه‌های میانگین گروهی.....
۱۴۹.....	۲-۷. استنتاج علی.....
۱۴۹.....	۱-۲-۷. مروری بر طرح‌های آزمایشگاهی با هدف ایجاد علیت.....
۱۵۴.....	۲-۲-۷. عدم قطعیت برآوردهای اثر علی ناشی از تصادفی‌سازی.....
۱۵۸.....	۳-۲-۷. استنتاج در انتخاب تصادفی آزمودنی‌های بکارگرفته‌شده تصادفی.....
۱۵۹.....	۳-۷. استنتاج بدون تصادفی‌سازی با نمونه‌گیری تصادفی.....
۱۶۰.....	۱-۳-۷. نمونه‌گیری تصادفی ساختگی.....
۱۶۳.....	۲-۳-۷. تصادفی‌سازی ساختگی.....
۱۶۶.....	۴-۷. خلاصه فصل: استنتاج علی متفاوت از استنتاج تعمیم‌دهنده است.....
۱۶۷.....	۵-۷. مطالعات پیشنهادی.....
۱۶۹.....	فصل هشت.....
۱۶۹.....	استنتاج بهتر در قرن بیست و یکم: جهانی فراتر از ($p < 0.05$).....
۱۷۶.....	منابع.....

فهرست جداول

- جدول ۱-۳: آماره‌های استنتاجی برای مثال شکاف دستمزد جنسیتی (اعداد آشکار شده) ۴۵
- جدول ۲-۳: توزیع تجمعی F برای چندک‌های منتخب Z براساس ۱۰۰۰۰ نمونه شبیه‌سازی شده ۴۸
- جدول ۱-۵: فراتحلیل برای ۲۰ مطالعه انفرادی براساس نمونه‌های تصادفی شبیه‌سازی شده ۸۵
- جدول ۲-۵: دو زمینه تکرار برای مطالعه اصلی شکاف دستمزد جنسیتی ۹۱
- جدول ۳-۵: حالت‌های مختلف عینیت و عدم‌عینیت معنادار برای تفاوت میانگین ۹۳
- جدول ۱-۶: قاعده تصمیم‌گیری بهینه و هزینه کل ۱۰۰۰۰ نمونه آزمون ۱۳۶
- جدول ۱-۷: حالت‌های استنتاجی در مقایسه‌های میانگین با توجه به نحوه تولید داده‌ها ۱۴۸
- جدول ۲-۷: تصادفی‌سازی‌های احتمالی برای دو گروه هم‌اندازه ۱۵۵

فهرست شکل‌ها

- شکل ۱-۲: مفهوم روایی ۲۳
- شکل ۱-۳: توزیع نمونه‌گیری تخمین شیب رگرسیون (β) و توزیع خطای نمونه‌گیری ($esam = \beta -$) ۳۸
- شکل ۲-۳: رویکرد کلی استنتاج آماری برای نمونه‌گیری تصادفی ساده ۴۴
- شکل ۳-۳: توزیع نسبت سیگنال به نویز (Z) براساس ۱۰۰۰۰ نمونه شبیه‌سازی شده ۴۹
- شکل ۱-۴: مروری کلی بر طرح‌های مختلف نمونه‌گیری ۵۵
- شکل ۱-۶: توزیع نمونه‌گیری برآورد شیب یا β (شکل بالا) و توزیع نسبت Z (شکل پایین) مشروط بر فرضیه تهی ۹۹
- شکل ۲-۶: مقدار احتمال ۵ درصد یک و دو طرفه منتج از توزیع نسبت سیگنال به نویز تحت فرضیه تهی ۱۰۲
- شکل ۳-۶: توزیع نسبت سیگنال به نویز (Z) با فرضیه تهی $H_0 = 0$ و تفاوت میانگین واقعی δ ۱۱۵
- شکل ۴-۶: حالت‌های مختلف «شکنجه داده‌ها» ۱۱۹
- شکل ۵-۶: بده-بستان میان اندازه خطای نوع اول (α) و اندازه خطای نوع دوم (β) ۱۳۱
- شکل ۶-۶: تعداد تصمیم‌های کاذب (ایجابی‌های کاذب و سلبی‌های کاذب) ۱۳۵