بررسی بین حجم نمونه با خطای نوع اول، دوم، توان و مقدار pآزمون

محمد غلامی[۱]

چکیده

استنباط آماری دارایس دو قسمت عمده، برآورد پارامترها و آزمون فرض ها می باشد. در این مقاله سعی داریم تا پس از بیان تعاریف مورد نیاز، به بررسی ارتباط خطای نوع اول، دوم و p- مقدار با حجم نمونه بپردازیم.

واژه های کلیدی: خطای نوع اول، خطای نوع دوم، توان آزمون، p-مقدار آزمون، حجم نمونه.

۱– مقدمه

استنباط آماری دارای دو قسمت عمده، برآورد پارامترها و آزمون فرض ها می باشد. هدف اصلی در برآورد پارامترها رسیدن به برآوردی حتی الامکان درست از پارامتر مجهول جامعه و در آزمون فرض ها، تعیین این موضوع که آیا با توجه به اطلاعات به دست آمده از نمونه، حدسی که درباره خصوصیتی از جامعه می زنیم به طور قوی تایید می شود یا خیر؟ از آنجایی که در تحلیل های آماری دسترسی به کل جامعه مقدور نمی‌باشد، برای برآورد پارامترها و استنباط در مورد پارامترها، ناگزیر از نمونه گیری با حجمی مشخص از جامعه می باشیم. از اینرو با توجه به اهمیت حجم نمونه در استنباط و برآورد آماری در این مقاله سعی داریم تا پس از بیان تعاریف مورد نیاز، به بررسی ارتباط خطای نوع اول، دوم و p-مقدار با حجم نمونه بپردازیم.

۲– تعاریف و مفاهیم اولیه از دیدگاه آماری[۲]

فرضیه آماری: فرضیه آماری ادعایی است در مورد توزیع احتمالی یک جامعه (که قابل نمونه گیری باشد)، اگر این فرضیه، توزیع احتمالی جامعه را کاملاً مشخص کند آن را فرضیه ساده و در غیر این صورت آن را فرضیه مرکب می گوئیم [۱].

فرضیه صفر (H_0): فرضیه صفر، فرضیه ای است که محقق درصدد تائید و یا رد آن می باشد.

فرضیه مقابل یا فرض یک (H_1): نفی ادعای فرضیه صفر را فرضیه مقابل می گوئیم.

خطای نوع اول (α): عبارتست از احتمال رد H_0 وقتیکه آن صحیح باشد.

خطای نوع دوم (β): عبارتست از احتمال قبول فرض H_0 وقتی که آن غلط باشد.

توان آزمون : عبارتست از احتمال رد فرضیه H_0 وقتیکه آن غلط است، در واقع توان آزمون برابر است با $1-\beta$ و عبارتست از توانایی ما در رد فرضیه غلط.

---

[۱] دانشجوی دکتری آمار زیستی دانشگاه تربیت مدرس

[۲] کلیه تعاریف استفاده شده در این بخش از کتاب آمار ریاضی دکتر پارسیان گرفته شده است

p- مقدار: کمترین مقداری از α است، که براساس یافته آماره آزمون ممکن است موجب رد فرضیه صفر گردد.
ناحیه بحرانی : مجموعه مقادیری که منجر به رد فرضیه صفر می شود را ناحیه بحرانی می گویند.
ناحیه پذیرش: مجموعه مقادیری از آماره آزمون که منجر به رد فرضیه یک می شود را ناحیه پذیرش می‌گویند .
تابع آزمون یا قاعده تصمیم گیری: تابع یا قاعده ای است که مجموعه همه مقادیر آماره آزمون را به دو ناحیه بحرانی و پذیرش تقسیم می کند [۱].
در جدول ۱ می توانید خلاصه تعاریف بالا را مشاهده کنید.

جدول ۱: نمایش خطای نوع اول، دوم، توان آزمون و سطح اطمینان بصورت جدولی

جامعه / نتیجه گیری	**فرضیه صفر درست است**	**فرضیه صفر غلط است**
رد فرضیه صفر	خطای نوع اول α	توان آزمون $1\text{-}\beta$
پذیرش فرضیه صفر	سطح اطمینان $1\text{-}\alpha$	خطای نوع دوم B

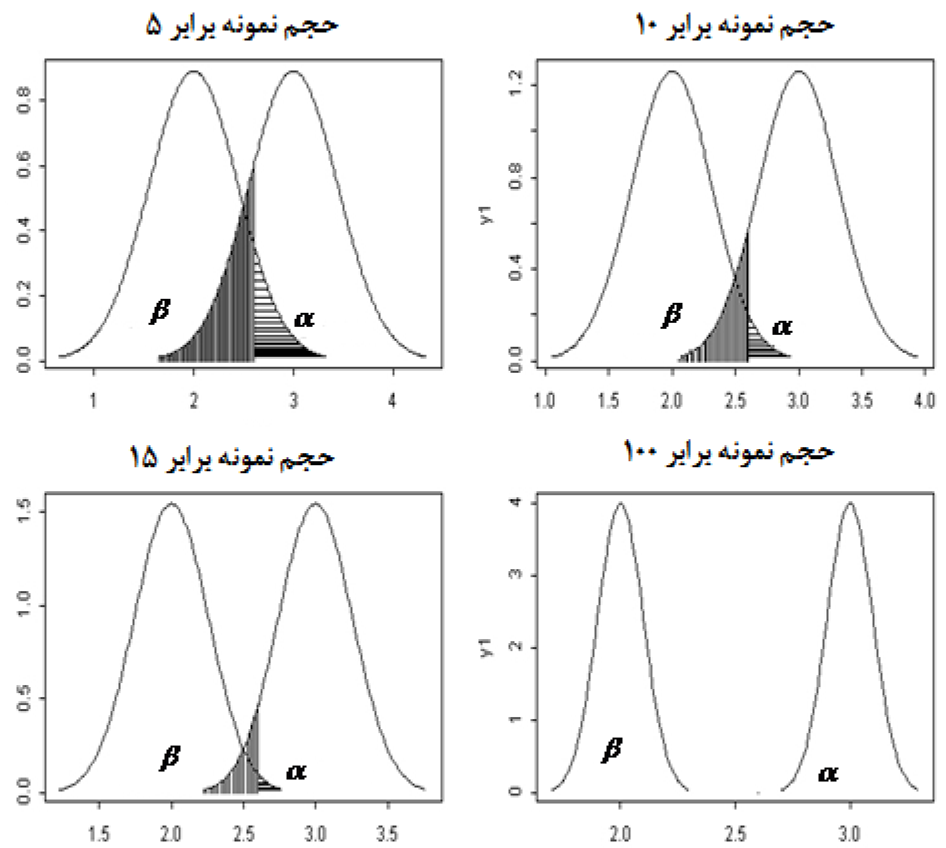
همانگونه که در جدول۱ مشخص است حاصل جمع توان آزمون و خطای نوع دوم و همچنین خطای نوع اول وسطح اطمینان برابر یک می باشد همچنین درآمار ریاضی (۱) ثابت می شود در صورتی که در آزمون فرضیه از لم نیمن پیرسن استفاده شود آنگاه $\alpha+\beta\leq 1$ می باشد و این بدین معنی است که α و β با یکدیگر رابطه معکوس دارند یعنی افزایش خطای دوم باعث کاهش نوع اول و بالعکس می شود.

مثال. فرض کنید مدت زمان بیهوشی توسط داروی خاص دارای توزیع طبیعی با میانگین μ و انحراف معیار ۱ باشد. یک متخصص بیهوشی نمونه ای تصادفی به حجم n از افرادی که توسط این دارو بیهوش کرده است انتخاب می کند، وی می خواهد آزمون $H_0: \mu=2 \text{ vs } H_1: \mu=3$ را انجام دهد. در این صورت خطای نوع اول در این آزمون این است که میانگین مدت زمان واقعی بیهوشی ۲ ساعت باشد و متخصص به اشتباه نتیجه گیری کند که این مدت زمان ۳ ساعت است. بلعکس خطای نوع دوم در این آزمایش تصمیم گیری در مورد ۲ ساعته بودن زمان بیهوشی در حالی که زمان واقعی ۳ ساعت است می باشد.

۳- ارتباط خطای نوع اول و دوم با حجم نمونه

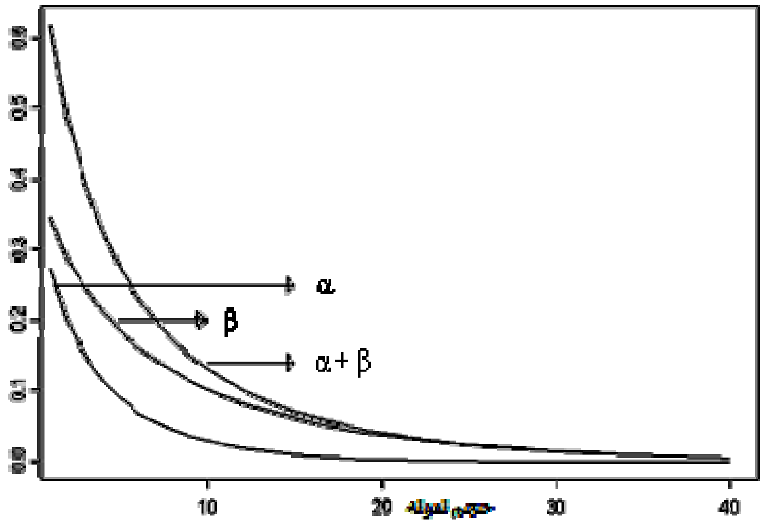
آنچه که ما را در برابر برآورد پارامتر های جامعه دچار عدم اطمینان و یا خطا می کند، این است که ما در برآورد پارامترهای جامعه از همه عناصر جامعه استفاده ننموده ایم. بدیهی است که انتظار داریم که هر چه تعداد نمونه ما در برآورد و آزمون فرضیه یک جامعه بیشتر می شود خطاها و عدم اطمینان ما درمورد پارامتر و فرضیه مورد ادعا کم می شود .در اینجا با ارائه یک مثال از آزمون فرضیه ساده از لحاظ ریاضی و همچنین گرافیکی ادعای بالا را ثابت می کنیم .اثبات کلی این مطلب برای هر نوع آزمون فرضیه ی خارج از سطح این مقاله می‌باشد. مثال متخصص بیهوشی را در نظر بگیرید. این متخصص برای رد فرضیه صفر خود ناحیه بحرانی $\bar{x} \geq ۲/۶$ را در نظر می گیرد براساس این ناحیه بحرانی α در این آزمون برابر $۱-\Phi(۰/۶\sqrt{n})$[1] و β برابر با $\Phi(-۰/۴n)$ می گردد. وقتی که حجم نمونه به بینهایت میل کند آنگاه $۰/۶\sqrt{n} \to \infty$ و $-۰/۴\sqrt{n} \to -\infty$، و در نتیجه این باعث صفر شدن α و β می گردد. این مطلب بخوبی در شکل ۱ نشان داده شده است.

---
[1] Φ نماد تابع احتمال تجمعی نرمال می باشد .



شکل ۱: نمایش گرافیکی مقادیر α و β

همان گونه که در شکل ۱ نمایش داده شده مساحت مناطق هاشور زده شده برابر مقادیر α و β می باشند که با افزایش حجم نمونه به سرعت به سمت صفر میل می کنند تا اینکه در $n=100$ تقریبا هیچ مساحتی برای محاسبه α و β مشاهده نمی شود که این خود نشان دهنده صفر شدن مقادیر α و β با افزایش حجم نمونه می‌باشد .این مطلب را نیز می توان به صورت دیگری نیز نمایش داد



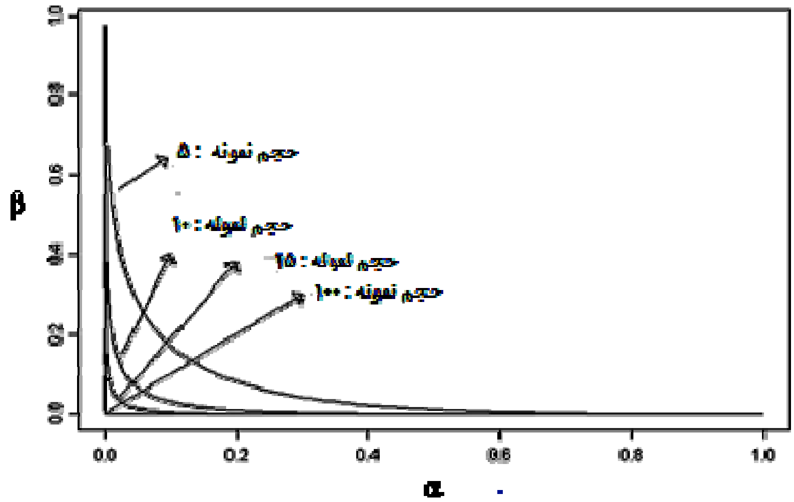
شکل ۲: نمایش α و β و مجموع آنها در مقابل حجم نمونه های مختلف

در شکل ۲ مقادیر α و β و همچنین مجموع آنها در مقابل حجم نمونه های متفاوت رسم شده است. در این شکل دو نکته به وضوح قابل مشاهده است:

۱- مجموع دو خطا کمتر از یک می باشد (براساس لم نیمن پیرسون).

۲- مقادیر α و β با افزایش حجم نمونه به سرعت به سمت صفر میل می کند.

در انتهای این بخش نظر شما را به شکل دیگری که می تواند بخوبی رابطه معکوس خطا ها و همچنین رابطه توام آنها را با حجم نمونه نشان دهد جلب می‌نمائیم.



شکل ۳: رابطه توام α و β با حجم نمونه

محور افقی شکل ۳، خطای α و محور عمودی آن مقادیر β می باشد. هنگامی که ما مرز ناحیه بحرانی را با توجه به یک حجم نمونه مشخص از منهای بینهایت تا مثبت بینهایت تغییر دهیم. مقادیر α و β براساس مقادیر مختلف ناحیه بحرانی تغییر می نمایند. مثلاً در مثال یک به ازای یک n مشخص مثلاً ۵ و k برابر با ۲/۶ ما به یک α و β مشخص دست یافتیم. حال اگر این k را تغییر دهیم به ازای هر k یک مقدار مشخص α و β بدست می آید.

شکل ۳، در واقع مجموع همه مقادیر α و β به ازای مقادیر مختلف k با توجه به n های متفاوت می باشد. این شکل نشان دهنده ۲ نکته می باشد:

۱- منحنی ها ی نزولی به ازای n های متفاوت، نشان دهنده این نکته است که به ازای یک n مشخص تغییرات α و β در عکس یکدیگر می باشد .

۲- افزایش شکم منحنی ها به سمت مرکز مختصات با افزایش n، نشان دهنده این است که α و β تواما با افزایش نمونه کاهش می یابند تا در حالت حدی (سرشماری جامعه) دقیقا بر محور مختصات (نبود خطای نوع اول و دوم) قرار گیرند.

۴- ارتباط p-مقدار با حجم نمونه

همانگونه که می دانیم p-مقدار به صورت

$$2*P(Z \geq |z_{obs}|)$$

تعریف می شود. که در آن $z_{obs} = \frac{(\bar{x}_n - \mu_\cdot)}{\sigma / \sqrt{n}}$ می باشد.

بنابراین در مثال متخصص بیهوشی p -مقدار برابراست با

$$p_{value} = ۲*(۱-\Phi(|\sqrt{n}\times(\bar{x}_n - ۲)|) \quad (۳)$$

از طرف دیگر با استفاده از قضیه قانون قوی اعداد بزرگ می دانیم (٤) که وقتی که n به سمت بینهایت میل می‌کند، میانگین نمونه ای به سمت میانگین واقعی جامعه میل میکند (میانگین واقعی جامعه را μ فرض کنید) و این بدین معنی است که $(\bar{x}_n - ۲) \approx (\mu - ۲)$ می باشد، با استفاده از این تقریب داریم.

$$\underset{n \to \infty}{|\sqrt{n} \times (\bar{x}_n - ۲)|} \approx \underset{n \to \infty}{|\sqrt{n} \times (\mu - ۲)|}$$

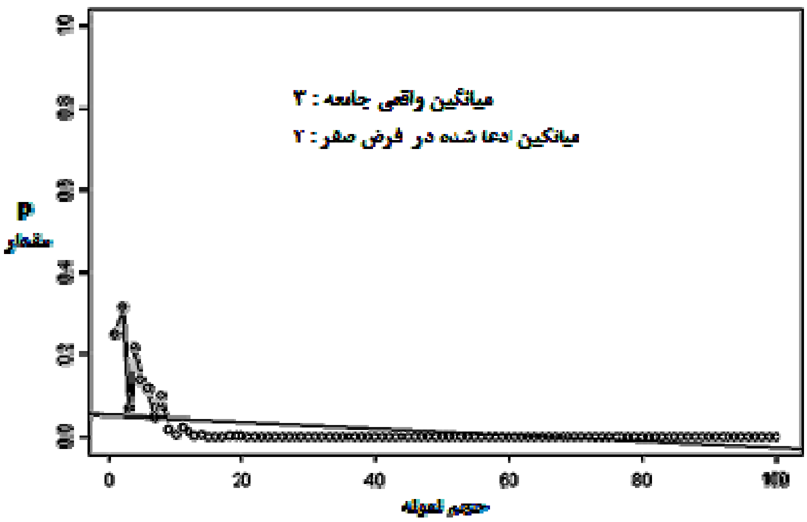
حال با در نظر گرفتن دو حالت درست و نادرست بودن فرضیه صفر به بررسی رابطه p-مقدار و حجم نمونه می‌پردازیم.

۴-۱- رابطه حجم نمونه و p -مقدار وقتی فرضیه صفر نادرست است

وقتی که پارامتر ادعا شده در فرضیه صفر مخالف پارامتر واقعی جامعه باشد.آنگاه $(\mu\text{-}2)$ به سمت c (c یک عدد ثابت میباشد) میل می کند و در نتیجه $\underset{n \to \infty}{|\sqrt{n} \times c| \to \infty}$ و از طرف دیگر با استفاده از خاصیت تابع توزیع نرمال داریم.

$$\Phi(\infty) = 1 \ \Rightarrow 1 - \Phi(\infty) = 0$$
$$p_{value} = 2 \times (1 - \Phi(|\sqrt{n} \times (\bar{x}_n - 2)|) \to 0$$

و این بدین معنی است که p-مقدار با افزایش نمونه کاهش می یابد. برای نشان دادن این موضوع، چون p-مقدار تابعی از حجم نمونه و میانگین نمونه ای می باشد برای بررسی رابطه حجم نمونه با p-مقدار به رسم شکل ٤ از طریق شبیه سازی پرداختیم برای رسم این شکل در ابتدا جامعه ای نرمال به حجم ۱۰۰۰ و میانگین سه و انحراف معیار یک تولید نمودیم. محور افقی شکل٤ اندازه نمونه و محور عمودی مقدار p-مقدار محاسبه شده می باشد. در محاسبه p-مقدار به منظور کاهش خطای نمونه گیری به جای یک بارنمونه گیری و محاسبه p-مقدار این عمل ۱۰ بار انجام شده و میانگین این ۱۰ بار جایگزین p-مقدار گذاشته شده است.

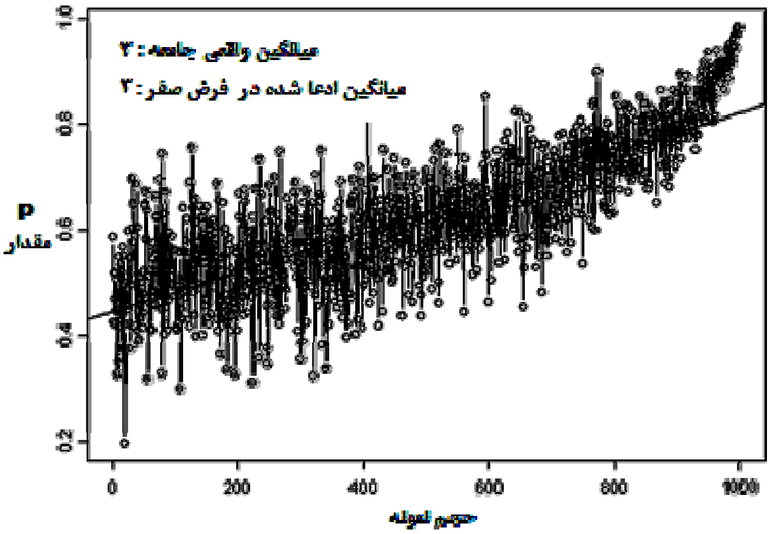


شکل ٤: رابطه حجم نمونه با pمقدار وقتی فرضیه صفر غلط است

همان گونه که انتظار داشتیم با افزایش حجم نمونه pمقدار به صورت نزولی به صفر همگرا می شود .توجیه این مطلب نیز بسیار ساده می باشدزیرا p-مقدار در واقع اعتبار فرضیه صفر است، وقتی که فرضیه صفر نادرست باشد p-مقدار ذاتا مقدار کوچکی(تقریبا صفر) است و افزایش نمونه در آشکار نمودن این مقدار کوچک نقش بسزایی را ایفا می کند، در واقع علت اینکه p-مقدار وقتی که n کوچک است بزرگ است اعتبار فرضیه صفر نیست، بلکه این بخاطر وجود خطای نمونه گیری در برآورد پارامتر میانگین می باشد .

۴-۲- رابطه حجم نمونه و p -مقدار وقتی فرضیه صفر درست است

وقتی که فرضیه صفر درست باشد با توجه به قانون قوی اعداد بزرگ میانگین نمونه ای به میانگین واقعی جامعه که درواقع میانگین ادعا شده در فرضیه صفر است میل می کند و این بدین معنی است که $(\bar{x}_n - ۲) \to ۰$ و از طرف دیگر با توجه به اینکه $\sqrt{n} \to \infty$ پاسخ $\lim\limits_{n \to \infty}(\sqrt{n} \times (\bar{x}_n - ۲))$ مبهم می باشد از طرفی چون میانگین نمونه ای مشخص نیست از لحاظ تحلیلی این حد قابل حل نمی باشد، برای حل این مشکل با استفاده از شبیه سازی آماری و تکنیک بیان شده در حالت نادرست بودن فرضیه صفر، شکل ۵ را برای مقادیر مختلف p-مقدار در برابر حجم نمونه های متفاوت رسم نمودیم .



شکل ۵: رابطه حجم نمونه با p-مقدار وقتی فرضیه صفر صحیح است

همان گونه که شکل ۵ نمایش می دهد با افزایش حجم نمونه p-مقدار آزمون به سمت یک میل می کند. توجیه این مطلب ساده است زیرا همانگونه که در قبل گفتیم p-مقدار درجه درستی فرضیه صفر را نشان می دهد وقتی که فرضیه صفر درست باشد p-مقدار ذاتا بزرگ است و افزایش حجم نمونه باعث آشکار شدن بیشتر این بزرگی می شود. علت اینکه p-مقدار وقتی که n کوچک است کوچک است کمی اعتبار فرضیه صفر نیست، بلکه این بخاطر وجود خطای نمونه گیری در برآورد پارامتر میانگین می باشد.

۵- بحث و نتیجه گیری

آنچه که ما را در برابر برآورد پارامتر های جامعه دچار عدم اطمینان و یا خطا می کند، این است که ما در برآورد پارامترهای جامعه از همه عناصر جامعه استفاده ننموده ایم. هنگامی که حجم نمونه افزایش می یابد α و β بصورت نزولی و توان آزمون و ضریب اطمینان بصورت افزایشی، صرفنظر از اینکه فرضیه صفر صحیح است و یا خیر افزایش می یابد. در مورد p-مقدار نیز در صورتی که فرضیه صفر صحیح باشد روند صعودی و در غیر این صورت روند نزولی را مشاهده می نمائیم.

منابع

[۱] محمد، ک.، ملک افضلی، ح. و نهاپتیان، و.،(۱۳۷۷). روش های آماری و شاخص های بهداشتی. نشر مولفین، تهران.

[۲] جباری نوقابی، ه. و جباری نوقابی، م. (۱۳۸۸). نکاتی چند در مورد برآورد حجم نمونه و معرفی نرم افزار مربوطه. نشریه دانشجویی آمار (ندا)، انجمن آمار ایران، شماره دوم، سال چهارم،۱۳-۲۱.

[۳] پارسیان، ا. (۱۳۸۰). مبانی آمار ریاضی. انتشارات دانشگاه صنعتی اصفهان، چاپ دوم، اصفهان.

[٤] بهبودیان، ج. (۱۳۸۱). روشهای ناپارامتری. انتشارات دانشگاه پیام نور، تهران.