مقاله بازآموزی

بر اساس تصویب دفتربازآموزی جامعه پزشکی وزارت‌بهداشت، درمان و آموزش پزشکی به پاسخ دهندگان پرسشهای مطرح شده در این مقاله ۱/۵ امتیاز به به کلیه مشمولین امتیاز بازآموزی تعلق می گیرد.

مقاله بازآموزی

آزمون فرضیه

نویسندگان: دکتر محمد فشارکی [۱] ـ دکتر عباس دانده پور [۲] ـ دکتر احمد عامری [۳]

مقدمه:

وقتی می خواهیم با استفاده از اطلاعات و مشاهده مربوط به نمونه در مورد جامعه تصمیمات مختلفی اتخاذ نمائیم چنین تصمیمی را تصمیم گیری آماری می نامند. معمولاً برای تصمیم گیری های آماری فرضهائی در جامعه آماری در نظر می گیرند.

فرضیه اظهار نظری است که درباره یک جمعیت یا چند جمعیت بیان می شود، اینگونه فرضیه ها در فرایند تحقیق به وسیله محقق مطرح می شود و پژوهشگر مایل است پس از جمع آوری داده ها به کمک روش های آماری درباره درست یا غلط بودن آن تصمیم گیری نماید.

فرضیه های آماری

در هر آزمون آماری، فرضیه اولیه ای وجود دارد که آن را فرضیه صفر (Null Hypothesis) می نامند و با علامت H_0 نشان می دهند. فرضیه صفر همیشه مساوی بودن دو موضوع و عدم اختلاف را مطرح می نماید.در اغلب موارد، فرضیه صفر همان فرضیه ای است که بخواهیم غلط بودن آن را ثابت کنیم.

در برابر این فرضیه، فرضیه های مخالف (Alternative Hypothesis) وجود دارند که آنها را با H_1 نشان می دهند.

خطای نوع اول و نوع دوم

خطای نوع اول یعنی غلط دانستن فرضیه H_0 در حالی که این فرضیه درست باشد (منفی کاذب) خطای نوع دوم یعنی قبول فرضیه H_0 در حالی که این فرضیه غلط باشد (مثبت کاذب) جدول زیر خطای نوع اول و خطای نوع دوم را نشان می دهد.

استنباط آماری درباره H_0	H_0 واقعا درست	H_0 واقعا غلط
نپذیرفتن H_0	خطای نوع اول α	استنباط درست
پذیرفتن H_0	استنباط درست	خطای نوع دوم β

آزمون یک دامنه و آزمون دو دامنه

اگر فرضیه صفر بر این باشد که تفاوتی بین میانگین ها و یا تفاوتی بین احتمال و عدم احتمال وجود ندارد، آزمون را یک دامنه می نامند و فرضیات آن به صورت:

۱و۳ـ عضو هیأت علمی دانشگاه علوم پزشکی ایران
۲ـ پزشک عمومی

Ho:P=q   یا Ho:μ1= μ2

H1:P=q   یا H1: μ1= μ2

می باشد. اگر فرضیه بر این باشد که میانگین یکی بیشتر یا کمتر از دیگری و یا تفاوت بین احتمال بیشتر و یا کمتر از عدم احتمال است، آزمون را یک دامنه می نامند و فرضیات آن به صورت زیر است:

Ho:P<q   یا Ho:μ1< μ2

H1:P>q   یا H1: μ1> μ2

آزمون z

الف: آزمون مقایسه میانگین با یک عدد فرضی μo هنگامی که δ مشخص باشد.

ملاک این آزمون فرمول ذیل است:

$$N>۳۰ \qquad z=\left|\frac{\bar{X}-\mu o}{\frac{\delta}{\sqrt{N}}}\right| \quad \text{ملاک}$$

X: میانگین نمونه   μo: میانگین فرضی

δ: انحراف معیار   N: اندازه جمعیت

حال اگر Z ملاک کوچکتر یا مساوی Z1-α/۲ در آزمون دو دامنه باشد، فرضیه صفر را قبول در غیر اینصورت آن را رد می نمائیم.

حال اگر Z ملاک کوچکتر یا مساوی Z1-α/۲ در آزمون یک دامنه باشد، فرضیه صفر را قبول در غیر اینصورت آن را در می‌نمائیم.

مثال: فرض کنید متوسط سطح کلسترول در کودکان یک منطقه ۱۷۵ mg/۱۰۰ ml، با انحراف معیار ۵۰ mg/۱۰۰ ml باشد، محققی ۴۹ کودک از این منطقه را به عنوان نمونه انتخاب می نماید و مشاهده می کند که متوسط سطح کلسترول برابر ۲۰۰ mg/۱۰۰ ml می باشد. این فرض که میانگین سطح کلسترول در این گروه بزرگتر از میانگین در کل جامعه است را آزمون کنید.

حل: فرضیه ها به صورت

Ho:μ1=175

H1:μ1> 175

خواهد بود:

$$z=\left|\frac{\bar{X}-\mu}{\frac{\delta}{\sqrt{N}}}\right| \qquad z=\left|\frac{۲۰۰-۱۷۵}{\frac{۵۰}{\sqrt{۴۹}}}\right|=۳/۵$$

اگر α = ۰/۰۵ فرض شود، عدد بحرانی از جدول سطح زیر منحنی نرمال ۱/۶۴ خواهد بود. چون مقدار ملاک آزمون از عدد بحرانی بزرگتر است، بنابراین فرضیه Ho را رد می نمائیم.

فرضیه ها به صورت:

در تست دو دامنه Ho:μ1= μ2 و H1:μ1= μ2

در تست یک دامنه Ho:μ1<μ2 و Ho:μ1> μ2

است ملاک آزمون به صورت:

$$z=\left|\frac{\bar{X}_1-\bar{X}_2}{\sqrt{\frac{\delta_1^2}{n_1}\ \frac{\delta_2^2}{n_2}}}\right|$$

می باشد.

X1: میانگین نمونه اول   X2: میانگین نمونه دوم

δ_1^2 و δ_2^2 به ترتیب واریانسهای دو جامعه هستند. روال آزمون عین قبل می باشد.

آزمون نسبت ها

با توجه به مطالعات قبلی برای محققی نسبت یا درصد جامعه مشخص می باشد و پژوهشگر می خواهد با نمونه ای که در اختیار دارد نسبت حاصل از این نمونه را با جامعه مقایسه نماید. دو حالت زیر را در نظر می گیریم:

I: برای مقایسه اختلاف نسبت صفت با یک نسبت فرضی

فرضیه ها به صورت:

در آزمون یک دامنه	در آزمون دو دامنه
Ho:K< Ko	Ho:K=Ko
H1:K>Ko	H1:K1=Ko

$$Z=\left|\frac{\frac{X}{n}-k_0}{\sqrt{\frac{K_0(1-k_0)}{n}}}\right|$$

K: نسبت صفت مورد مطالعه در جامعه

Ko: نسبت مشخص شده

X: تعداد افراد واحد صفت مورد مطالعه

N: تعداد نمونه

حال اگر $|Z| < z1\text{-}\alpha/2$ در یک آزمون دو دامنه باشد، فرضیه صفر را قبول در غیر اینصورت آن را رد می کنم.

حال اگر $|Z| < z1\text{-}\alpha$ در آزمون یک دامنه باشد فرضیه صفر را قبول در غیر اینصورت آن را رد می نمائیم.

II: برای مقایسه نسبت دو صفت فرضیه ها به صورت:

$Ho: K1 \geq K2$	$Ho: K1 = K2$
در آزمون یک دامنه	در آزمون دو دامنه
$H1: K1 < K2$	$H1: K1 \neq K2$

$$Z=\left|\frac{n_1 \quad n_2}{\sqrt{K^{\wedge}(1-K^{\wedge})(\frac{1}{n_1}+\frac{1}{n_2})}}\right|$$

K1 و K2 : به ترتیب معرف نسبت صفت مورد مطالعه

n1 و n2: حجم نمونه اول و دوم

X1 و X2: به ترتیب فرد مورد مطالعه

$$\hat{k}=\frac{x_1+x_2}{n_1+n_2}$$

و بقیه روند آزمون طبق گذشته می باشد.

آزمون t

در این قسمت مواردی بیان می گردد که در آن N<30 باشد، در این حالت منحنی نرمال در وسط برآمدگی شدیدی پیدا می کند و در گوشه ها مساحت زیادتری را در بر می گیرد، چنین توزیعی را توزیع تی می نامند.

جدول توزیع تی شامل دو مقدار یک ضریب خط و دیگری درجه آزادی (Degree of Freedom) می باشد که آن را با df نمایش می دهد.

df = تعداد قیود ـ تعداد مشاهدات

الف: مقایسه میانگین با عدد ثابت وقتی که تی معلوم نباشد.

فرضیه ها به صورت:

$Ho: \mu < \mu o$	$Ho: \mu = \mu o$
در آزمون یک دامنه	در آزمون دو دامنه
$H1: \mu > \mu o$	$H1: \mu < \mu o$

است. ملاک آزمون به صورت:

$$t=\left|\frac{\bar{X}-\mu}{\frac{S}{\sqrt{N}}}\right|$$

خواهد بود، که تحت فرضیه Ho این ملاک دارای توزیع t با df=n-1 خواهد بود حال اگر:

در آزمون دو دامنه $t < t1\text{-}\alpha/2$ df ملاک

یا در آزمون یک دامنه $t < t1\text{-}\alpha$ df ملاک باشد، فرضیه صفر را قبول، در غیر اینصورت آن را رد می نمائیم.

ب: آزمون مقایسه اختلاف میانگین دو جامعه ولی واریانس دو جامعه معلوم نباشد.

فرضیه ها به صورت:

$Ho: \mu 1 < \mu 2$	$Ho: \mu 1 = \mu 2$
در آزمون یک دامنه	در آزمون دو دامنه
$H1: \mu 1 > \mu 2$	$H1: \mu 1 < \mu 2$

سه حالت زیر را در نظر می گیریم:

۱ـ اگر حجم دو نمونه برابر نباشند.

ملاک آزمون به صورت:

$$t=\left|\frac{\bar{X}_1-\bar{X}_2}{\sqrt{\left[\frac{\sum x_1^2+\sum x_2^2}{N_1+N_2-2}\right]\left[\frac{1}{N_1}+\frac{1}{N_2}\right]}}\right|$$

و

$$\sum x_I^2 = \sum x_I^2 - \frac{(\sum x_I)}{N}$$

که تحت فرضیه Ho ملاک t محاسبه شده دارای توزیع نرمال با df=N1+N2-2 خواهد بود. حال اگر در آزمون دو دامنه $t < t1\text{-}\alpha/2, df$ یا در آزمون یک دامنه $t < t1\text{-}\alpha , df$ باشد. فرضیه صفر را قبول در غیر اینصورت آن را رد می نمائیم.

مثال: محققی می خواهد ثابت کند که Po2 در دو گروه افراد سیگاری و غیر سیگاری متفاوت می باشد. بدین منظور ۹ نفر از افراد سیگاری و ۹ نفر از افراد غیر سیگاری را مورد مطالعه قرار داده و میزان Po2 در دو گروه در جدول زیر ثبت می نماید.

Po2 افراد غیرسیگاری مرد (۴۵-۲۵) ساله	۵۳ ۴۸ ۳۶ ۳۲ ۳۸ ۳۴ ۲۶ ۴۳ ۱۵
Po2 افراد سیگاری مرد (۴۵-۲۵) ساله	۲۴ ۳۱ ۲۹ ۳۹ ۳۱ ۵۲ ۴۳ ۳۹ ۵۱

آیا اختلاف مشاهده شده بین دو گروه اختلاف تصادفی بوده

حل: فرضیه به صورت

Ho: $\mu1 = \mu2$

H1: $\mu1 = \mu2$

X_1=۳۶/۱۱

X_2=۳۸/۷۷

$$\sum X_۱^۲ = ۱۲۷۸۳ - \frac{(۳۲۵)^۲}{۹} = ۱۰۴۶/۸۸$$

$$\sum X_۲^۲ = ۱۴۲۱۵ - \frac{(۳۴۹)^۲}{۹} = ۶۸۱/۵۵$$

$$t = \left| \frac{۳۶/۱۱ - ۳۸/۸}{\sqrt{(\frac{۱۰۴۶/۸۸ + ۶۸۱/۵۵}{۹+۹-۲})(\frac{۱}{۹} + \frac{۱}{۹})}} \right| = ۰/۵۵$$

df=n۱+n۲-۲=۱۶

اگر α ۰/۰۵ = فرض شود از جـدول ۲/۱۲ = ۱۶ و ۰/۹۷۵ t. پس فرضیه صفر قبول می شود یعنی اختلاف مشـاهده بیـن هر گروه از نظر آماری معنی دار نمی باشد.

۲ـ اگر حجم دو نمونه برابر باشد.

اگر n۱=n۲ باشد فرمول ملاک به صورت زیر خواهد بود:

$$t = \left| \frac{\bar{X}1 - \bar{X}2}{\sqrt{\frac{\sum x_1^2 + \sum x_2^2}{N(N-1)}}} \right|$$

و df=2(N-1) و بقیه روال آزمون عین قبل خواهد بود.

۳ـ اگر میانگین دو گروه مستقل نبوده و نوعی همبستگی بین آنها وجود داشته باشد از آزمون مقایسه زوج ها (Paired t-Test) به صورت زیر استفاده می گردد.

$$t \text{ ملاک} = \left| \frac{Xd}{\sqrt{\frac{\Sigma x^۲ d}{N(N-1)}}} \right|$$

کـــــــــه در آن: df=N-1 $\quad Xd = X_1 - X_2$

$$\sum X^2 d = \sum X_d^2 - \frac{(\sum \times d)^2}{N} \qquad X_d = \frac{\sum Xd}{N}$$

حال اگر df و $t<t_1-\alpha/2$ مـلاک در یـک آزمـون دو دامنـه باشـد فرضیـه صفـر را قبـول در غـــیر اینصـــورت آن را رد می نمائیم. اگر $t<t1\text{-}\alpha, df$ ملاک در یک آزمون یک دامنـه باشد. فرضیـه صفـر را قبـول در غـیر ایـن صـورت آن را رد می نمائیم.

ج ـ مواردی که از آزمون t نمی توان استفاده کرد.

در صورتی آزمون t به صورتهائی که بیان شد مـورد استفاده قرار می گیرد، که واریانس گروهها یکسان باشند، بـه عبـارت

دیگر داده ها از یک جامعه آماری باشند. در صورتی که واریانس گروهها یکسان نباشند از فرمول t به صورت ذیل استفاده می گردد.

$$t = \left| \frac{X1 - X2}{\sqrt{\frac{\Sigma x_1^2 +}{N_1(N_1-1)} + \frac{\Sigma x_2^2}{N_2(N_2-1)}}} \right|$$

فرمول بالا دارای دو درجه آزادی خواهد بود یکی df1 و دیگری df2 و $df_1=n_1-1$ و $df_2=n_2-1$ خال اگر t ملاک از هر دو t جدول بزرگتر باشد فرضیه صفر را رد می کنیم، چنانچه t حساب شده از هر دو t جدول (یا مساوی) کوچکتر باشد، فرضیه صفر را قبول نموده، اگر t ملاک از یک t جدول بزرگتر و از t دیگر کوچکتر باشد آن را با t متوسط کد فرمول آن به صورت زیر است مقایسه می نمائیم.

$$t = \frac{(N_1-1)t_1+(N_2-1)t_2}{N_1+N_2-2}$$

References:

1- Bernard Rosner, ***Fundamentals of Biostatistic*** Harvard University, 1993.

2- David E. Matthews/Vernon T. Fare Well, ***Using and undereatanding Medical statistics*** Karger, 1988.

3- Michael J. Campbell/ David Machin, ***Medical Statistics. A Common sense Approach***, University of Sheffled. UK. 1999.

4- Johannes, I.Polly, F.Bancroft,s ***Introductoin to biostatists.*** Harper, Row, New york. Evanston, and london 1974.

۵ـ آیت اللهی ، محمد تقی (ترجمه) ـ ***اصول و روشهای آمار زیستی***، انتشارات امیر کبیر ۱۳۶۸.

۶ـ فشارکی، محمد ـ ***آمار زیستی*** ، انتشارات عبادی فر ۱۳۸۰.

سوالات مقاله بازآموزی آزمون فرضیه

۱ـ در مقایسه میانگین دو جامعه با انجام یک آزمون مناسب فرضیه $Ho{:}\mu1=\mu2$ در سطح خطای ۵ درصد رد نشده است مفهوم آن چیست؟ (جامع علوم پایه - شهریور ۷۴)

الف: در ۹۵٪ از موارد میانگین نمونه جامعه اول با جامعه دوم اختلاف دارد.

ب: بر مبنای یافته های مذکور اطمینان کافی برای رد فرضیه Ho وجود دارند.

ج: با اطمینان بیش از ۹۵٪ میانگین دو جامعه مساوی است.

د: خطای نوع دوم حداکثر برابر ۵٪ است.

۲ـ در مقایسه دو میانگین هنگامی از ملاک t استفاده می کنیم که: (جامع علوم پایه ـ اسفند ۷۳)

الف: واریانس دو جامعه معلوم و مساوی باشند.

ب: واریانس دو جامعه معلوم و نامساوی باشند.

ج: واریانس دو جامعه نامعلوم و نامساوی باشند.

د: واریانس دو جامعه نامعلوم و مساوی باشند.

۳ـ در آزمونی که فرضیه Ho درست است به طور متوسط از هر ۸۰ آزمون چه تعداد آن در سطح اشتباه ۵٪ رد می شود؟ (جامع علوم پایه ـ شهریور ۷۶)

الف: ۲ ب: ۴ ج:۶ د:۸

۴ـ کدام شرط از شرایط زیر برای انجام آزمون مساوی بودن میانگین در دو گروه (t-test) باید برقرار باشد؟ (جامع علوم پایه ـ شهریور ۷۸).

الف: تعداد مشاهدات دو گروه باید مساوی باشند.

ب: اطلاعات باید از جوامع نرمال انتخاب شده باشد.

ج: میانگین دو گروه باید تقریباً مساوی باشد.

د: نمونه ها باید کوچک باشند.

۵ـ در یک مطالعه، اوره خون تعدادی بیمار اندازه گیری شده و در مرحله بعد بیماران تحت درمان قرار گرفته و پس از پایان دوره درمان اوره آنها مجدداً اندازه گیری گردید. فرضیه محقق عبارت است از «میانگین اوره خون بعد از درمان مورد نظر نسبت به قبل از آن به طور معنی داری تغییر می کند» در جهت رد فرضیه مزبور،کدامیک از آزمون های زیر مناسب تر است؟ (دستیاری ـ اسفند ۷۷).

الف: آزمون t زوجی

ب: آزمون t

ج: آزمون مقایسه نسبت ها

د: آزمون های غیر پارامتری

۶ـ یک روش درمان برای بیمارانی که از یک نوع بیمار قلبی رنج می برند موجب بهبود ۷۰ درصد انها شد، ۱۰۰ نفر از مبتلایان به همین بیماری با روش دیگر تحت درمان قرار گرفتند که ۸۵ نفر آنها بهبود یافتند، آیا احتمال بهبود یافتن بیماران با روش دوم نیز ۷۰ درصد است؟

الف: بله ب: خیر

۷ـ برای ازمون فرضیه $Ho{:}\mu1=\mu2$ که در ان $\mu1$ متوسط تعداد دندانهای پوسیده در بچه های دبستانی در شمال شهر و $\mu2$ متوسط تعداد دندانهای پوسیده بچه های دبستانی در جنوب شهر است. از چه آزمونی استفاده کنیم؟

الف: آزمون t
ب: آزمون مقایسه زوج ها
ج: آزمون Z (برای نسبت)
د: آزمونZ (مقایسه دو صفت)

۸ـ در ازمون فرضیه $Ho{:}\mu = \mu_0$ اندازه احتمال برای اشتباه نوع اول عبارت است از احتمال: (جامع علوم پایه ـ اسفند ۷۵)

الف: رد فرضیه مخالف به شرط صحیح بودن آن
ب: رد فرضیه Ho به شرط صحیح بودن آن
ج: رد فرضیه Ho
د: رد فرضیه مخالف

۹ـ محققی سطح قابل قبول برای رد فرضیه آماری (صفر) را از ۵ درصد تغییر داده است این تصمیم موجب کدام مورد می شود؟

الف: افزایش خطای نوع اول
ب: کاهش خطای نوع اول
ج: افزایش توان آزمون آماری
د:خطای نوع اول تغییر نمی کند

۱۰ـ در ازمون فرضیه اگر میانگین دو گروه مورد بررسی مستقل نبوده و نوعی همبستگی بین انها وجود داشته باشد از چه ازمونی استفاده می شود؟

الف:ازمون Z
ب:آزمون T
ج:آزمون مقایسه زوج ها (Paired T Test)
د: هیچکدام