



بِه نام خدا

آمار در علوم اجتماعی

رشته علوم اجتماعی

۲ واحد درسی

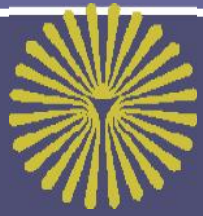
نام منبع و مؤلف : آمار در علوم اجتماعی (از فصل دوم به بعد)، دکتر کریم منصور فر انتشارات دانشگاه پیام نور، ۱۳۷۴

Payam Noor University Ebook

PNUeb

علی محمد جوادی، دستیار آموزشی دانشگاه پیام
نور مرکز گیلانغرب

۱



طرح درس

فصل دوم : لزوم وارد شدن نظریه احتمال در روشهای آماری

فصل سوم : متغیر تصادفی

فصل چهارم : برآورد و اصول تخمین

فصل پنجم : آزمونهای آماری

Payam Noor University Ebook

PNUeb

علی محمد جوادی ، دستیار آموزشی دانشگاه پیام
نور مرکز گیلانغرب



جایگاه درس

درس آمار در علوم اجتماعی از دروس پایه دوره کارشناسی

علوم اجتماعی است.

Payam Noor University Ebook

PNUeb

علی محمد جوادی، دستیار آموزشی دانشگاه پیام
نور مرکز گیلانغرب



عناوین فصل اول

- ❖ فضای نمونه یا فضای پیشامد ساده
- ❖ تعریف احتمال
- ❖ تعریف احتمال بر مبنای فراوانی نسبی
- ❖ قضایای مربوط به احتمال
- ❖ احتمال هندسی
- ❖ احتمال شرطی



عناوین فصل اوّل

- ❖ تعریف پیشامدهای مستقل و نامستقل
- ❖ قضایای مربوط به حاصل جمع و حاصل ضرب دو پیشامد مستقل
- ❖ قضایای مربوط به حاصل جمع و حاصل ضرب دو پیشامد وابسته
- ❖ قضیه بیس
- ❖ حل مسایل احتمالات بوسیله دیاگرام درخت
- ❖ آزمایشهای تکراری

Payam Noor University Ebook

PNUeb

علی محمد جوادی، دستیار آموزشی دانشگاه پیام
نور مرکز گیلانغرب

۵



فضای نمونه یا فضای پشامد

فضای نمونه یا فضای حوادث ساده :

مجموعه‌ای که عناصر آن ، نمایش تمام نتایج ممکنه در یک آزمایش باشد فضای نمونه نام دارد و آن را با علامت Ω و یا S نشان می دهند . (آزمایش عملی است که نتیجه آن معلوم نباشد)



فضای نمونه یا فضای حوادث ساده

نکته : اگر K سکه را همزمان و یا یک سکه را K بار پرتاب کنیم تعداد حالات ممکنه برابر با 2^K

پس اگر K تاس همزمان و یا K تاس را K بار پرتاب داد حالات ممکنه برابر خواهد بود با 6^K



فضای نمونه یا فضای حوادث ساده

❖ اگر کیسه ای دارای n مهره باشد تعداد حالات ممکنه برای اینکه m ($m < n$) مهره انتخاب شود برابر است با :

$$C_n^m = \frac{n!}{(n-m)!m!}$$



پیشامد یا حادثه

هر عضو از یک فضای نمونه را یک پیشامد یا حادثه گویند (هر زیرمجموعه‌ای از فضای نمونه را یک حادثه گوئیم و با حروف بزرگ A, B, \dots نمایش می‌دهیم).



انواع پیشامد

❖ پیشامد حتمی یا یقینی

❖ پیشامد غیر ممکن

❖ پیشامد تصادفی

Payam Noor University Ebook

PNUeb

علی محمد جوادی، دستیار آموزشی دانشگاه پیام
نور مرکز گیلانغرب

۱۰



پیشامد حتمی یا یقینی

پیشامدی که تحت هر شرایط به طور اجتناب ناپذیر رخ دهد ،
پیشامد حتمی نام دارد و آنرا با علامت | نشان می دهند.
مثلاً در ریختن یک تاس معمولی آمدن رویه کمتر از ۷ یک
پیشامد حتمی است.



پیشامد غیر ممکن

پیشامدی که رخ دادن آن تحت هیچ شرایطی هرگز ممکن نباشد ،
پیشامد غیر ممکن نام دارد و آنرا با O نمایش می دهیم. مثلاً در
ریختن یک تاس معمولی آمدن عدد بزرگتر از ۶ غیر ممکن
است.



پیشامد تصادفی

پیشامدی که ممکن است وقوع یابد یا وقوع نیابد مانند آمدن رویه ۵ در یک بار پرتاب تاس پیشامد تصادفی نام دارد.



تعریف احتمال

احتمال پیشامد A عددی است که اندازه امکان وقوع را نشان دهد. اگر یک آزمایش برای هر N حالت مختلف، نتایج محتمل یکسان به دست دهد، و اگر n حالت ($n < N$) برای پیشامد A که با $p(A)$ نشان داده می شود عبارت است از:

$$p(A) = \frac{n}{N}$$

و یا



تعريف احتمال

اگر نتایج یک آزمایش بتواند کلاً به N حالت هم احتمال (یعنی از لحاظ وقوع پیشامد هیچ گونه امتیازی به هم نداشته باشند) و ناسازگار (مانعت الجمع یعنی با وقوع یکی از آنها وقوع حالات دیگر امکانپذیر نباشد) واقع شود و n حالت آن برای پیشامد معین A مساعد باشد احتمال وقوع پیشامد A کسری است برابر:

$$p(A) = \frac{n}{N}$$



تعريف احتمال

به عبارت ساده تر
نسبت تعداد حالات مساعد بر تعداد حالات ممکنه را احتمال می
نامند

$$\text{احتمال} = \frac{\text{تعداد حالات مساعد برای حادثه } A}{\text{تعداد حالات ممکنه}}$$



تابع احتمال

قاعده یا قانون تناظری را گویند که با هر پیشامد A در فضای نمونه یک عدد حقیقی $p(A)$ را که احتمال پیشامد A نامیده میشود، مربوط کند به طوری که برای هر پیشامد A ،

$$\text{اولاً: } p(A) \geq 0$$

ثانیاً: مجموع احتمالات مربوط به کلیه پیشامدهای متمایز مساوی یک گردد.

ثالثاً اگر پیشامدهای A, B ناسازگار باشند آنگاه تساوی زیر صادق باشد

$$p(A \text{ یا } B) = p(A) + p(B)$$

این خاصیت قضیه مجموع احتمالات نامیده می شود.



تعریف احتمال بر مبنای فراوانی نسبی

سکه ای را n بار می اندازیم اگر تعداد شیر آمدن آنرا n_A بنامیم آنگاه فراوانی نسبی شیر آمدن سکه برابر $\frac{n_A}{n}$ خواهد بود حال اگر تعداد آزمایش را زیادتر کنیم این فراوانی نسبی به عدد $\frac{1}{2}$ نزدیک می شود. یعنی برای مقادیر بزرگ n یک سری پیشامد های تصادفی نسبتاً ثابت می ماند که این مقدار ثابت را اندازه امکان وقوع می نامند و به عنوان مقدار تقریبی احتمال پیشامد تصادفی

$$p(A) = \frac{n_A}{n} \text{ قبول می شود.}$$



احتمال بر مبنای فراوانی نسبی

نکته: در عمل ، احتمال همان فراوانی نسبی است که برای بیشترین تعداد آزمایش به دست آمده باشد.



قضایای مربوط به احتمال

احتمال پیشامد غیر ممکن صفر است. $p(O)=0$

احتمال پیشامد یقینی مساوی یک است. $p(I)=1$

برای هر پیشامد دلخواه A عددی وجود دارد که بین صفر و یک است:
 $0 \leq p(A) \leq 1$



قضایای مربوط به احتمال

۴. اگر پیشامد A زیر مجموعه پیشامد B باشد ($A \subset B$) آنگاه رابطه $p(A) \leq p(B)$ برقرار خواهد بود.

۵. اگر پیشامد های A, B هم ارز باشند ($A=B$) آنگاه احتمالهای آنها مساوی خواهند بود.
 $p(A)=p(B)$

۶. مجموع احتمال وقوع پیشامد A و عدم وقوع پیشامد A یعنی (\bar{A}) مساوی است با یک.

$$p(A) + p(\bar{A}) = 1$$



فضای نمونه یا فضای حوادث ساده

❖ اگر کیسه ای دارای n مهره باشد تعداد حالات ممکنه برای اینکه m ($m < n$) مهره انتخاب شود برابر است با :

$$C_n^m = \frac{n!}{(n-m)!m!}$$



قضایای مربوط به احتمال

۷. قضیه حاصل جمع احتمالات - اگر پیشامد A به S حالت ناسازگار A_1, A_2, \dots, A_s تجزیه گردد، یعنی:

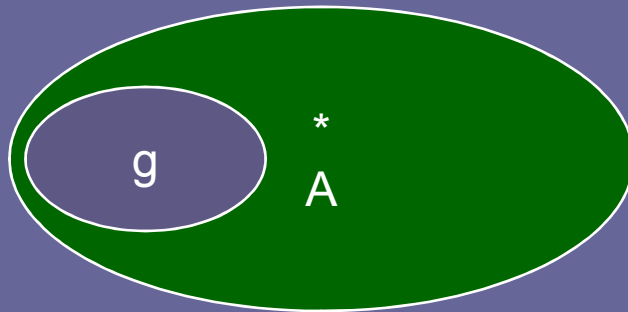
$$A = A_1 + A_2 + \dots + A_s, \quad A = A_1 \cup A_2 \cup \dots \cup A_s$$

آنگاه احتمال پیشامد A که آن را پیشامد مرکب می نامند مساوی خواهد بود با مجموع احتمالهای تک تک آنها.

$$p(A) = p(A_1) + p(A_2) + \dots + p(A_s) = \sum_{i=1}^s p(A_i)$$



احتمال هندسی



احتمال اینکه نقطه A در درون ناحیه g باشد برابر است با نسبت وسعت اندازه G بر وسعت اندازه G یعنی:

اندازه وسعت ناحیه g

$$P(A \text{ در } g) = \frac{\text{اندازه وسعت ناحیه } g}{\text{اندازه وسعت ناحیه } G}$$

وسعت ناحیه G



احتمال هندسی

$$p(A \in g) = \frac{mes.g}{mes.G}$$

در احتمال هندسی نیز خواص زیر برقرار است :

$$0 \leq p(A \in g) \leq 1 \quad ; \quad p(\emptyset) = 0 \quad ; \quad p(A \in G) = 1$$



تعریف : حوادث شرطی

اگر رخ دادن یک حادثه (مانند A) مشروط به چگونگی وقوع حادثه دیگر (مانند B) باشد، این دو حادثه را نسبت به هم شرطی می گویند بدیهی است که حوادث شرطی نسبت به هم وابسته اند.



احتمال شرطی

احتمال وقوع پیشامد A هنگامی که پیشامد B قبلاً اتفاق افتاده باشد احتمال شرطی نامیده می شود و به صورت $p(A|B)$ نشان داده میشود و چنین خوانده می شود: احتمال وقوع پیشامد A به شرط آنکه پیشامد B قبلاً وقوع یافته باشد.



احتمال شرطی

در نظریه احتمال مطلب زیر را به عنوان تعریف احتمال شرطی قبول کرده اند:

به شرط آنکه $p(B) \neq 0$ باشد.

$$p(A|B) = \frac{p(A \cdot B)}{p(B)}$$
$$\left(p(A|B) = \frac{p(A \cap B)}{p(B)} \right)$$



پیشامدهای مستقل

دو پیشامد A, B را مستقل از هم نامند که وقوع یکی روی وقوع دیگری تأثیر نداشته باشد به عبارت دیگر احتمال وقوع پیشامد با احتمال شرطی آن پیشامد یکسان باشد، یعنی:

$$p(A|B) = p(A) \quad \text{همچنین} \quad p(B|A) = p(B)$$

اگر پیشامد A مستقل از پیشامد B باشد پیشامد B نیز مستقل از پیشامد A خواهد بود.



پیشامد های مستقل

نکته: هرگاه از جامعه نمونه ای برداریم و دوباره آن نمونه را به جای خود باز گردانیم یعنی عمل جاگذاری را انجام دهیم آنگاه پیشامد اول تأثیری بر احتمال پیشامد دوم نخواهد داشت.



پیشامدهای نامستقل (وابسته)

دو پیشامد A, B را وابسته به هم نامند هرگاه وقوع یکی روی دیگری تأثیر داشته باشد.



پیشامدهای نامستقل (وابسته)

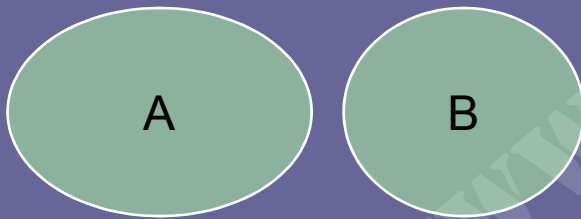
نکته: اگر از یک جامعه نمونه ای برداریم و دوباره آنرا در جای خود قرار ندهیم یعنی بدون جایگذاری نمونه گیری انجام دهیم آنگاه این دو پیشامد را وابسته به هم می نامند.



قضیه حاصل جمع دو پیشامد ناسازگار

اگر دو پیشامد A, B ناسازگار باشند (مانند شکل زیر) آنگاه احتمال حاصل جمع این دو پیشامد برابر با حاصل جمع تک تک آنها خواهد بود. یعنی:

$$p(A \cup B) = p(A) + p(B)$$



احتمال مذکور برای S پیشامد مستقل برابر است با:

$$p\left(\sum_{i=1}^s A_i\right) = \sum_{i=1}^s p(A_i)$$

علی محمد جوادی، دستیار آموزشی دانشگاه پیام
نور مرکز گیلانغرب



قضیه حاصل ضرب دو پیشامد مستقل

اگر دو پیشامد A, B مستقل از هم باشند آنگاه احتمال حاصل ضرب این دو پیشامد مساوی است با حاصل ضرب احتمال های آن دو پیشامد.

$$p(A \cap B) = p(A) \cdot p(B)$$

این قضیه را برای S پیشامد مستقل A_1, A_2, \dots, A_s نیز می توان

تعمیم داد

$$p(A_1 \cap A_2 \cap \dots \cap A_s) = p(A_1) \cdot p(A_2) \cdot \dots \cdot p(A_s)$$



حاصل جمع دو پیشامد در حالت کلی

اگر دو پیشامد A, B داشته باشیم در این صورت احتمال حاصل جمع این دو پیشامد برابر با حاصل جمع احتمال تک تک آنها منهای احتمال اشتراک دو پیشامد، که ممکن است این احتمال صفر

$$p(A \cup B) = p(A) + p(B) - p(A \cap B) \quad \text{باشد}$$



قضیه حاصل جمع سه پیشامد در حالت کلی

قضیه حاصل جمع را در مورد S پیشامد نیز می توان به کار برد .
مثلاً برای سه پیشامد A, B, C عبارت است از :

$$p(A \cup B \cup C) = p(A) + p(B) + p(C) - p(A \cap B) - p(A \cap C) - p(B \cap C) + p(A \cap B \cap C)$$

ملاحظه می شود که پیشامد های فرد دارای علامت مثبت و
پیشامدهای زوج دارای علامت منفی است.



قضیه حاصل ضرب دو پیشامد وابسته

احتمال حاصل ضرب دو پیشامد وابسته مساوی است با حاصل ضرب احتمال یکی از این پیشامدها در احتمال شرطی پیشامد دیگر به شرطی که پیشامد قبل وقوع یافته باشد. یعنی:

$$p(A \cap B) = p(A) \cdot p(B|A) = p(B) \cdot p(A|B)$$

این فرمول از طرفین و وسطین کردن احتمال شرطی بدست آمده است.



قضیه حاصل ضرب دو پیشامد وابسته

قضیه حاصل ضرب دو پیشامد وابسته را می توان برای S پیشامد A_1, A_2, \dots, A_s تعمیم داد.

$$p(A_1 \cap A_2 \cap \dots \cap A_s) = p(A_1) \cdot p(A_2 | A_1) \dots p(A_s | A_1 A_2 \dots A_{s-1})$$

این قضیه به نام قضیه عمومی حاصل ضرب احتمالها نامیده می شود.

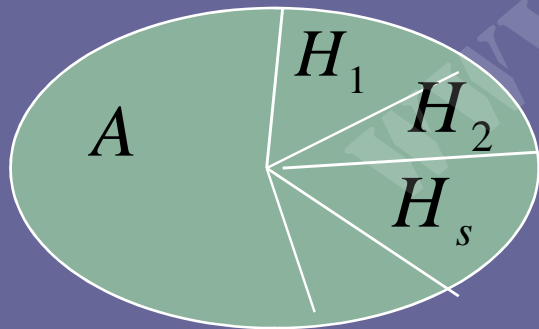


قضیه بیس (بیز)

اگر H_1, H_2, \dots, H_s پیشامدهایی ناسازگار باشند که یکی از آنها باید در یک آزمایش رخ دهد، یعنی:

$$p(H_1) + p(H_2) + \dots + p(H_s) = 1$$

و A عبارت از پیشامدی که برای آن $p(A) \neq 0$ باشد.



Payam Noor University Ebook

علی محمد جوادی، دستیار آموزشی دانشگاه پیام
نور مرکز گیلانغرب



قضیه بیس (بیز)

در اینصورت احتمال شرطی $p(H_i|A)$ برای هریک از پیشامدهای H_1, H_2, \dots, H_s به شرطی که پیشامد A رخ داده باشد از فرمول زیر محاسبه می شود:

$$p(H_i|A) = \frac{p(H_i) \times p(A|H_i)}{\sum_{i=1}^s p(H_i) \cdot p(A|H_i)}$$

مخرج کسر را احتمال متوسط می نامند و آنرا با $p(A)$ نشان می دهند.



درخت

اگر تعداد آزمایشها (n) محدود باشد، می توان شمارش پیشامدهای ممکن و مساعد را با استفاده از یک روش ترسیمی که به نام دیاگرام درخت نامیده می شود به راحتی تعیین و احتمال های این قبیل پیشامدها را به آسانی محاسبه کرد



آزمایشهای تکراری

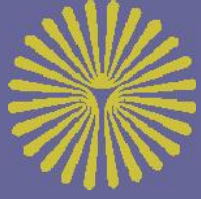
اگر در یک آزمایش تکراری p احتمال موفقیت یک پیشامد در یک آزمایش و q احتمال عدم وقوع باشد. احتمال اینکه در n آزمایش مورد نظر موفقیت آن آزمایش درست m بار تکرار شود توسط فرمول زیر بیان می شود:

$$p_n(m) = C_n^m p^m q^{n-m}$$

که در آن $p_n(m)$ نشانگر احتمالی است که در n بار آزمایش

پیشامد A به تعداد m بار رخ می دهد و

$$C_n^m = \frac{n!}{m!(n-m)!}$$



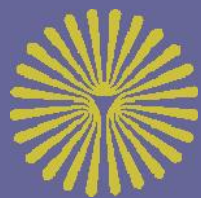
عناوین فصل دوم

- ❖ متغیر تصادفی
- ❖ متغیر تصادفی گسسته
- ❖ متغیر تصادفی پیوسته
- ❖ تابع توزیع
- ❖ تابع احتمال
- ❖ قانون اعداد بزرگ
- ❖ قضیه حد مرکزی

Payam Noor University Ebook

PNUeb

علی محمد جوادی، دستیار آموزشی دانشگاه پیام
نور مرکز گیلانغرب



عناوین فصل دوم

- ❖ امید ریاضی یک متغیر تصادفی
- ❖ خواص امید ریاضی
- ❖ امید ریاضی یک متغیر تصادفی پیوسته
- ❖ واریانس و انحراف معیار متغیر تصادفی گسسته
- ❖ خواص واریانس
- ❖ توزیع های معیار برای متغیر گسسته
- ❖ توزیع های معیار برای متغیر پیوسته

Payam Noor University Ebook

PNUeb

علی محمد جوادی، دستیار آموزشی دانشگاه پیام
نور مرکز گیلانغرب



متغیر تصادفی

متغیر هایی را که بر حسب نتیجه آزمایش مقدار اختیار می کنند متغیر تصادفی می نامند

Payam Noor University Ebook

PNUeb

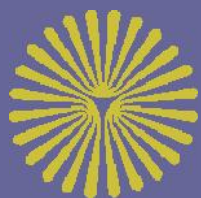
علی محمد جوادی، دستیار آموزشی دانشگاه پیام
نور مرکز گیلانغرب

۴۵



ناحيه تعريف

مجموعه مقادیری که متغیر تصادفی می تواند اختیار کند مجموعه مقادیر ممکن متغیر یا ناحیه تعريف آن نامیده می شود



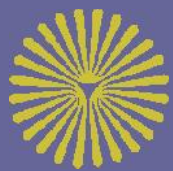
نکته

در ریاضیات اگر رابطه ای میان متغیر X, Y داشتیم برای X مقادیر را خودمان انتخاب می کردیم و این انتخاب تصادفی نبود بلکه کاملاً انتسابی بود در صورتیکه انتخاب متغیرهای تصادفی اساساً اختیاری نیست در هر آزمایش مقدار متفاوتی به دست می آید که مقدار آن کاملاً تصادفی است و قبل از آزمایش نمی توان گفت که چه مقداری برای نتیجه آزمایش حاصل خواهد شد.



قرارداد

متغیرهای تصادفی را با حروف بزرگ مانند X, Y, \dots و مقادیر
متغیرها را با حروف کوچک مانند x, y, \dots نشان خواهیم داد.



متغیر تصادفی گسسته

متغیر تصادفی را گسسته نامند که فقط بتواند مقادیر معینی را در روی خط اعداد گویا اختیار کند. به عبارتی متغیر تصادفی که بتواند مجموعه اعداد شمارش پذیر را اختیار کند متغیر تصادفی تصادفی گسسته نامیده می شود .



انواع متغیر تصادفی گسسته

- . متغیر تصادفی با توزیع یکنواخت
- . متغیر تصادفی با توزیع دو جمله ای
- . متغیر تصادفی با توزیع پواسن

Payam Noor University Ebook

PNUeb

علی محمد جوادی، دستیار آموزشی دانشگاه پیام
نور مرکز گیلانغرب

۵۰

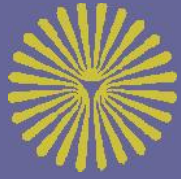


متغیر تصادفی با توزیع یکنواخت

متغیری است که به ازاء تمامی مقادیر آن احتمالهایش ثابت بماند

یعنی:

$$p(x) = \frac{1}{n}$$



متغیر تصادفی با توزیع دو جمله ای

متغیر تصادفی X که میدان تغییرات آن مجموعه اعداد صحیح مثبت از صفر تا n تشکیل می دهد و احتمال متناظر با آن مقادیر توسط جمله عمومی خیام نیوتن (فرمول برنولی) بیان می شود متغیر تصادفی با توزیع دو جمله ای نامیده می شود:

$$p_x(x) = C_n^x p^x q^{n-x}$$



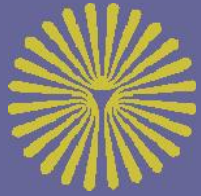
عناوین فصل دوم

- ❖ متغیر تصادفی
- ❖ متغیر تصادفی گسسته
- ❖ متغیر تصادفی پیوسته
- ❖ تابع توزیع
- ❖ تابع احتمال
- ❖ قانون اعداد بزرگ
- ❖ قضیه حد مرکزی

Payam Noor University Ebook

PNUeb

علی محمد جوادی، دستیار آموزشی دانشگاه پیام
نور مرکز گیلانغرب



عناوین فصل دوم

- ❖ امید ریاضی یک متغیر تصادفی
- ❖ خواص امید ریاضی
- ❖ امید ریاضی یک متغیر تصادفی پیوسته
- ❖ واریانس و انحراف معیار متغیر تصادفی گسسته
- ❖ خواص واریانس
- ❖ توزیع های معیار برای متغیر گسسته
- ❖ توزیع های معیار برای متغیر پیوسته

Payam Noor University Ebook

PNUeb

علی محمد جوادی، دستیار آموزشی دانشگاه پیام
نور مرکز گیلانغرب



متغیر تصادفی

متغیر هایی را که بر حسب نتیجه آزمایش مقدار اختیار می کنند متغیر تصادفی می نامند

Payam Noor University Ebook

PNUeb

علی محمد جوادی، دستیار آموزشی دانشگاه پیام نور مرکز گیلانغرب

۵۵



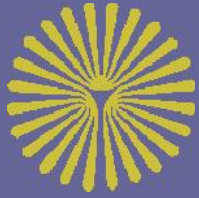
ناحيه تعريف

مجموعه مقادیری که متغیر تصادفی می تواند اختیار کند مجموعه مقادیر ممکن متغیر یا ناحیه تعريف آن نامیده می شود



نکته

در ریاضیات اگر رابطه ای میان متغیر X, Y داشتیم برای X مقادیر را خودمان انتخاب می کردیم و این انتخاب تصادفی نبود بلکه کاملاً انتسابی بود در صورتیکه انتخاب متغیرهای تصادفی اساساً اختیاری نیست در هر آزمایش مقدار متفاوتی به دست می آید که مقدار آن کاملاً تصادفی است و قبل از آزمایش نمی توان گفت که چه مقداری برای نتیجه آزمایش حاصل خواهد شد.



قرارداد

متغیرهای تصادفی را با حروف بزرگ مانند X, Y, \dots و مقادیر
متغیرها را با حروف کوچک مانند x, y, \dots نشان خواهیم داد.



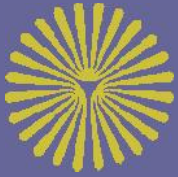
متغیر تصادفی گسسته

متغیر تصادفی را گسسته نامند که فقط بتواند مقادیر معینی را در روی خط اعداد گویا اختیار کند. به عبارتی متغیر تصادفی که بتواند مجموعه اعداد شمارش پذیر را اختیار کند متغیر تصادفی تصادفی گسسته نامیده می شود .



انواع متغیر تصادفی گسسته

- . متغیر تصادفی با توزیع یکنواخت
- . متغیر تصادفی با توزیع دو جمله ای
- . متغیر تصادفی با توزیع پواسن



متغیر تصادفی با توزیع یکنواخت

متغیری است که به ازاء تمامی مقادیر آن احتمالهايش ثابت بماند

یعنی:

$$p(x) = \frac{1}{n}$$



متغیر تصادفی با توزیع دو جمله ای

متغیر تصادفی X که میدان تغییرات آن مجموعه اعداد صحیح مثبت از صفر تا n تشکیل می دهد و احتمال متناظر با آن مقادیر توسط جمله عمومی خیام نیوتن (فرمول برنولی) بیان می شود متغیر تصادفی با توزیع دو جمله ای نامیده می شود:

$$p_x(x) = C_n^x p^x q^{n-x}$$



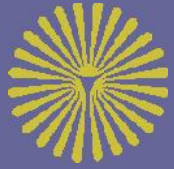
متغیر تصادفی با توزیع پواسن

متغیر تصادفی ناپیوسته X که میدان تغییرات آن مجموعه اعداد غیر منفی باشد و احتمالهای مقادیر X با فرمول

$$p(x) = \frac{\{^x \cdot e^{-}\}}{x!}$$

بیان شده باشد آن را متغیر تصادفی پواسن می نامند.

در فرمول تقریبی پواسن $\} = np$ (میانگین) مقدار ثابت است و e پایه دستگاه لگاریتمهای طبیعی است و مقدار آن تقریباً $۷۲/۲$ می باشد.



نکات

نکته: هنگامی از فرمول پواسن استفاده می شود که تعداد زیادی وقایع مستقل از هم صورت می گیرد ولی برای هر یک از آنها احتمال کوچکی وجود دارد که پیشامد معینی اتفاق افتد. (وقایع کمیاب)

نکته: در توزیع پواسن واریانس و میانگین با هم برابرند، یعنی:

$$u^2 = \}$$

Payam Noor University Ebook

PNUeb

علی محمد جوادی، دستیار آموزشی دانشگاه پیام نور مرکز گیلانغرب



متغیر تصادفی پیوسته

متغیر تصادفی پیوسته X ، تمامی مقادیر ممکن واقع در یک فاصله
رامی تواند قبول کند. بنابراین برای متغیر تصادفی پیوسته باید
احتمال فاصله‌ها را در نظر بگیریم ..



تابع چگالی

در توزیعهای پیوسته احتمال اینکه X دقیقاً یکی از مقادیر را اختیار کند برابر صفر است در نتیجه امکان نوشتن تابع احتمال به صورت جدول نیست بلکه تابع را فقط می توان به صورت فرمول نوشت.

توزیع احتمالات X را بوسیله نمایش تابع $f(x)$ در نظر می گیریم و $f(x)$ را تابع چگالی توزیع احتمالات و یا به طور ساده تر تابع چگالی می نامیم.



تابع چگالی

تابع چگالی برای یک متغیر تصادفی پیوسته X تابعی است مانند $f(x)$ با خواص زیر: $f(x) \geq 0$

سطح زیر منحنی برابر با یک است؛

$$\int_{-\infty}^{+\infty} f(x) dx = 1$$

احتمال اینکه X در فاصله بین a و b باشد

$$\int_a^b f(x) dx = p(a < x < b)$$



تعريف

اگر متغير تصادفئ پیوسته X دارای تابع چگالی احتمال $f(x)$ باشد آنگاه تابع توزیع تجمعی X به صورت زیر تعریف می شود

$$F(x) = \int_{-\infty}^x f(x) dx$$

پس: $F'(x) = f(x)$ (منظور از $F'(x)$ مشتق $F(x)$ است.)



تابع توزیع

فرض کنید X یک متغیر تصادفی و x یک عدد حقیقی دلخواه باشد. احتمال اینکه متغیر تصادفی X مقداری کوچکتر یا مساوی x اختیار کند تابع توزیع احتمال متغیر تصادفی X گفته می‌شود و آن را با $F_x(x)$ نشان می‌دهند بنابراین مفهوم تابع توزیع به صورت زیر نشان داده می‌شود:

$$F_x(x) = p\{X \leq x\}$$



تابع احتمال

جدول یا فرمولی که تمام مقادیر متغیر تصادفی گسسته را همراه با احتمالهای متناظرشان نشان دهد تابع احتمال نامیده می‌شود و معمولاً آن را با $p(x)$ و یا با $g(x)$ نشان می‌دهند.

$$p(x) = p(X = x)$$

برای متغیر تصادفی گسسته گاهی به جای تابع احتمال آن را توزیع احتمال می‌گویند.



فرمول کلی تابع احتمال

فرمول کلی تابع احتمال برابر است با :

$$p(x) = C_n^x p^x q^{n-x}$$



قانون فوق هندسی

کیسه ای دارای N_p مهره سفید و N_q مهره سیاه است یعنی $N = N_p + N_q$ نمونه ای مرکب از n مهره استخراج می کنیم $n < N$ متغیر تصادفی X عبارتست از تعداد مهره های سفید در این نمونه. تابع احتمال عبارت است از:

$$p(x) = \frac{C_{N_p}^x \cdot C_{N_q}^{n-x}}{C_N^n}$$

و آن را قانون فوق هندسی می گویند.



قانون اعداد بزرگ

قانون اعداد بزرگ ارتباط نزدیک احتمال یک پیشامد را با فراوانی نسبی آن در یک سری آزمایش که تعداد آنها به اندازه کافی زیاد باشد برقرار می‌کند. یعنی هرچه تعداد آزمایش زیادتر گردد فراوانی نسبی نیز به سمت احتمال حقیقی وقوع همان پیشامد میل خواهد کرد.



قانون اعداد بزرگ

قانون اعداد بزرگ شامل قضیه برنولی و قضیه پواسون است.

Payam Noor University Ebook

PNUeb

علی محمد جوادی، دستیار آموزشی دانشگاه پیام
نور مرکز گیلانغرب

۷۴



قضیه برنولی

اگر احتمال وقوع پیشامد معین A در کلیه آزمایشها ثابت بماند، یعنی ثابت $p(A)$ آنگاه با افزایش نامحدود تعداد آزمایشها (n) فراوانی نسبی $\left(\frac{n_A}{n}\right)$ به احتمال وقوع آن پیشامد نزدیک و نزدیکتر می‌گردد. یعنی:

$$\lim_{n \rightarrow \infty} \frac{n_A}{n} = p(A)$$



قضیه پواسون

اگر احتمال وقوع پیشامد A در n آزمایش برابر با p_1, p_2, \dots, p_n باشد به طوری که $p_M = \frac{p_1 + p_2 + \dots + p_n}{n}$ گردد آنگاه با احتمال نزدیک به یک، فراوانی نسبی این پیشامد در صورتی که تعداد آزمایش به اندازه کافی زیاد باشد به احتمال متوسط (p_M) وقوع آن نزدیک خواهد شد. یعنی:

$$\left| \frac{n_A}{n} - p_M \right| < \nu$$



قضیه حد مرکزی

اگر به صورت تصادفی از یک جامعه نامحدود نمونه ای با حجم $n \geq 30$ انتخاب شود:

- میانگین نمونه دارای توزیع نرمال است.
- اندازه میانگین این توزیع با میانگین جامعه برابر است.
- این توزیع دارای انحراف معیاری است که به خطای استاندارد یا خطای معیار میانگین معروف است و با t_d و یا t_x نمایش می دهند یعنی $t_d = \frac{t}{\sqrt{n}}$ که در آن n حجم نمونه و t انحراف معیار جامعه است.

Payam Noor University Ebook

PNUEB

علی محمد جوادی، دستیار آموزشی دانشگاه پیام نور مرکز گیلانغرب



نکته

اگر جامعه میانگینها را استاندارد کنیم یعنی از هر میانگین $(\bar{x}_i) \sim$ ،
را کسر کرده و به انحراف معیار میانگینها تقسیم نماییم. به این

صورت:

$$U_i = \frac{\bar{x}_i - \bar{x}}{\frac{\sigma}{\sqrt{n}}}$$

آنگاه این متغیر جدید U دارای میانگین صفر و واریانس یک $\lim_{n \rightarrow \infty}$ خواهد بود. یعنی:

$$U \sim N(0,1) \text{ (توزیع استاندارد جامعه میانگینها)}$$



نکته

طبق قضیه حد مرکزی اگر توزیع متغیر مورد مطالعه نرمال باشد توزیع میانگین هم نرمال خواهد بود



امید ریاضی یک متغیر تصادفی

میانگین حسابی تمامی مقادیر ممکن کمیت تصادفی در تئوری احتمالات امید ریاضی نامیده می شود.

امید ریاضی متغیر تصادفی عددی است که نشان می دهد به طور متوسط چه مقداری از متغیر تصادفی را در آزمایش باید انتظار داشت.



امید ریاضی (متغیر گسسته)

$$x_s, \dots, x_2, x_1$$

مقادیر ممکن یک متغیر تصادفی X بوده

$$p_s, \dots, p_2, p_1$$

نیز

احتمالهای متناظر این مقادیر باشند به طوری که $\sum_{i=1}^s p_i = 1$

باشد در این صورت $E(X) = p_1 x_1 + p_2 x_2 + \dots + p_s x_s$

را امید ریاضی متغیر تصادفی X می نامند و آن را به صورت زیر نشان می دهند:

$$E(X) = \sum_{i=1}^s p_i x_i$$

که در آن p_s, \dots, p_2, p_1 وزن متغیر تصادفی نامیده می شوند.



نکته

اگر متغیر تصادفی X مقادیر به تعداد نامحدود شمارش پذیر
را با احتمالهای $p_1, p_2, \dots, p_i, \dots$ اختیار کند
آنگاه امید ریاضی آن کمیت عبارت است از

$$E(X) = \sum_{i=1}^{\infty} p_i x_i$$



تعریف (امید ریاضی)

اگر p احتمال وقوع یک رویداد در یک آزمایش و X تعداد رویدادها در n آزمایش تکراری باشد، تعداد رویدادهای مورد انتظار یعنی امید ریاضی آن پیشامد در n آزمایش برابر است با:

$$E(X) = np$$



خواص امید ریاضی

۱- اگر متغیر تصادفی همیشه مقدار ثابت C را اختیار کند آنگاه امید ریاضی این متغیر تصادفی همان مقدار ثابت C خواهد بود .

$$E(c) = c$$



خواص امید ریاضی

۲- اگر X را متغیر تصادفی و C را عدد حقیقی در نظر بگیریم
آنگاه:

$$E(cX) = cE(X) \quad \text{و} \quad E(X + c) = E(X) + c$$



خواص امید ریاضی

۳- اگر متغیر تصادفی X بتواند مقادیر X_1, X_2, \dots, X_n و متغیر تصادفی Y بتواند مقادیر Y_1, Y_2, \dots, Y_n را قبول کند امید ریاضی متغیر تصادفی $X + Y$ که از ترکیب دو متغیر تشکیل شده است مساوی است با حاصل جمع دو امید ریاضی یعنی:

$$E(X + Y) = E(X) + E(Y)$$

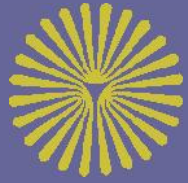
قاعده فوق را جمع پذیر بودن امید ریاضی می نامند.



خواص امید ریاضی

۴- امید ریاضی حاصلضرب دو متغیر تصادفی مستقل از یکدیگر برابر است با حاصلضرب امید ریاضی آن دو متغیر تصادفی یعنی اگر دو متغیر تصادفی X, Y از هم مستقل باشند داریم:

$$E(XY) = E(X).E(Y)$$



نکته

دو خاصیت اخیر را می توان برای n متغیر تصادفی X_n, \dots, X_2, X_1 نتیجه گرفت:

$$E(X_1 + X_2 + \dots + X_n) = E(X_1) + E(X_2) + \dots + E(X_n)$$

و همچنین

$$E(X_1 X_2 \dots X_n) = E(X_1) \times E(X_2) \times \dots \times E(X_n)$$



امید ریاضی یک متغیر تصادفی پیوسته

نکته: اگر متغیر تصادفی X پیوسته باشد، تابع احتمال را تابع چگالی احتمال می‌نامند.

$f(x)$

اگر متغیر تصادفی X پیوسته بوده و دارای تابع چگالی احتمال باشد یعنی:

$$\int_{-\infty}^{+\infty} f(x)dx = 1$$

آنگاه امید ریاضی متغیر X بوسیله انتگرال زیر تعریف می‌شود:

$$E(x) = \int_{-\infty}^{+\infty} x.f(x)dx$$

البته به شرط آنکه انتگرال فوق دارای جواب باشد.



واریانس متغیر تصادفی گسسته

واریانس متغیر تصادفی عبارتست از امید ریاضی توان دوم انحراف متغیر تصادفی X از امید ریاضی خود:

$$D(X) = E[X - E(X)]^2$$

$$D(X) = E(X^2) - [E(X)]^2$$

و یا

یعنی واریانس متغیر تصادفی برابر است با امید ریاضی مجذور متغیر تصادفی یعنی $E(X^2)$ منهای مجذور امید ریاضی متغیر تصادفی X .



انحراف معيار متغير تصادفی

انحراف معيار متغير تصادفی عبارت است از مجذور مثبت واریانس و آن را با u نشان می دهند یعنی:

$$u = \sqrt{D(X)}$$



خواص واریانس

۱- اگر متغیر تصادفی X ثابت باشد یعنی $X=C$ در اینصورت:

$$D(X) = D(c) = 0$$



خواص واریانس

۲- اگر متغیر تصادفی X به مقدار ثابت c تقسیم شود واریانس کل برابر است با :

$$D\left(\frac{X}{c}\right) = \frac{D(X)}{c^2}$$

و اگر ضرب کنیم :

$$D(cX) = c^2 [D(X)]$$



خواص واریانس

۳- حاصل جمع (یا تفاضل) واریانس دو متغیر تصادفی مستقل و ناسازگار X, Y عبارتست از حاصل جمع (یا تفاضل) تک تک متغیرهای تصادفی X, Y یعنی:

$$D(X \pm Y) = D(X) \pm D(Y)$$



توزیع‌های معیاربرای متغیر گسسته

۱- توزیع یکنواخت:

امید ریاضی و واریانس متغیر تصادفی با توزیع یکنواخت به قرار زیر است:

$$D(X) = \overline{x^2} - (\overline{x})^2 \quad \text{و} \quad E(X) = \overline{x}$$

$$\overline{x^2} = \frac{x_1^2 + x_2^2 + \dots + x_n^2}{n} \quad \text{و} \quad \overline{x} = \frac{x_1 + x_2 + \dots + x_n}{n} \quad \text{که:}$$



توزیع های معیار برای متغیر گسسته

۲- توزیع دو جمله ای:

$$b(x, n, p) \text{ با پارامترهای } n, p \text{ و نماد } p(x) = C_n^x p^x q^{n-x}$$

توزیعی نامتقارن است تنها در صورتی متقارن می باشد که

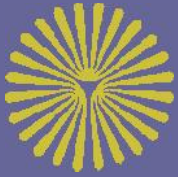
$$p = \frac{1}{2} \text{ گردد.}$$

امید ریاضی و واریانس توزیع دو جمله ای عبارت است از:

$$E(x) = np \quad p(x) = C_n^x p^x q^{n-x}$$

Payam Noor University Ebook

علی محمد جوادی، دستیار آموزشی دانشگاه پیام
نور مرکز گیلانغرب



توزیع های معیار برای متغیر گسسته

۱- توزیع پواسون:

امید ریاضی و واریانس توزیع پواسن عبارت است از:
 $p(x) = \frac{e^{-\lambda} \lambda^x}{x!}$ که در آن λ پارامتری مثبت است. امید

$$E(x) = \lambda \quad \text{و} \quad D(x) = \lambda$$

امید ریاضی و واریانس این توزیع با هم برابرند.



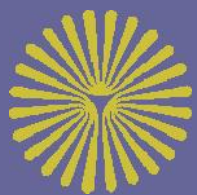
تعریف چگالی برای متغیر پیوسته

تابع چگالی برای یک متغیر تصادفی پیوسته X تابعی است مانند $f(x)$ با خواص زیر:

$$f(x) \geq 0$$

$$\int_{-\infty}^{\infty} f(x) dx = 1$$

$$\int_a^b f(x) dx = p(a < x < b)$$



توزیع‌های معیار برای متغیر پیوسته

۱- توزیع یکنواخت: تابع چگالی توزیع یکنواخت در فاصله $f(x) = c$ معین (a, b) ثابت است و در خارج آن صفر می‌باشد.

$$\text{که در آن } c = \frac{1}{b-a} \text{ پس } a \leq x \leq b \quad f(x) = \begin{cases} \frac{1}{b-a} \\ 0 \end{cases}$$

در سایر نقاط



توزیع‌های معیار برای متغیر پیوسته

امید ریاضی و واریانس توزیع یکنواخت در فاصله $[a, b]$ عبارت است از:

$$D(x) = \frac{1}{12}(b - a)^2 \quad \text{و} \quad E(X) = \frac{1}{2}(b + a)$$

این توزیع در مواردی نظیر، مطالعه گرد کردن اعداد و طول زمان استفاده می‌شود



توزیع‌های معیار برای متغیر پیوسته

۲- توزیع نرمال: تابع چگالی منحنی نرمال عبارتست از:

$$f(x) = \frac{1}{u\sqrt{2f}} e^{-\frac{(x-\sim)^2}{2u^2}}$$

که در آن $e = 2/72$ و $f = 3/14$

و نیز $-\infty < x < +\infty$ و $-\infty < \sim < +\infty$

$$0 < u^2 < +\infty$$



توزیع‌های معیار برای متغیر پیوسته

با تغییر متغیر $u = \frac{x - \bar{x}}{u}$ تابع چگالی منحنی نرمال به صورت زیر درمی‌آید:

$$y = \frac{1}{u \sqrt{2f}} e^{-\frac{u^2}{2}}$$

در این صورت آنرا منحنی نرمال استاندارد شده می‌گویند که دارای میانگین صفر و انحراف معیار یک است.



توزیع‌های معیار برای متغیر پیوسته

نمودار نرمال

WWW.PNUeB.COM

Payam Noor University Ebook

PNUeB

علی محمد جوادی، دستیار آموزشی دانشگاه پیام
نور مرکز گیلانغرب

۱۰۳



توزیع‌های معیار برای متغیر پیوسته

خواص منحنی نرمال :

❖ منحنی فوق متقارن است یعنی میانگین، میانه و نما با هم برابر است.

$$Q_3 = -Q_1 = 0.6745$$

❖ چارکهای Q_3 و Q_1 عبارتند از:

❖ میدان تغییرات صفت در منحنی نرمال $(R = x_{\max} - x_{\min})$ تقریباً ۶ برابر انحراف معیار است.



توزیع‌های معیار برای متغیر پیوسته

❖ میانگین قدر مطلق انحرافات توزیع نرمال برابر است:

$$" = u \sqrt{\frac{2}{f}} = 0.7979 \times u$$

❖ نسبت انحراف معیار بر میانگین قدر مطلق انحرافات تقریباً برابر

با ۱/۲۵ می باشد

$$\frac{s}{" } \# 1 / 25$$

"



توزیع‌های معیار برای متغیر پیوسته

❖ ضرایب چولگی و کشیدگی آن صفر است ($A = k = 0$) شکل پراکندگی منحنی نرمال به مقدار انحراف معیار u مربوط است اگر u کوچک باشد پراکندگی کمتر و اگر u بزرگ باشد پراکندگی توزیع جامعه از نرمال ($u = 1$) بیشتر است.



توزیع‌های معیار برای متغیر پیوسته

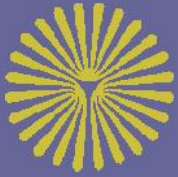
۳- متغیر تصادفی پیوسته با توزیع نمایی: متغیر تصادفی پیوسته X که مجموعه مقادیر ممکن آن تمامی اعداد حقیقی غیر منفی در فاصله $(0, +\infty)$ باشد و چگالی آن را در این فاصله با فرمول $f(x) = \dots$ بیان شده باشد و در آن $\{$ عددی ثابت و مثبت است و آن را پارامتر توزیع نمایی می‌نامند.

امید ریاضی و واریانس آن برابر است با :

$$D(x) = \frac{1}{\lambda^2}$$

$$E(X) = \frac{1}{\lambda}$$

علی محمد جوادی، دستیار آموزشی دانشگاه پیام نور مرکز گیلانغرب



توزیع‌های معیار برای متغیر پیوسته

۴- توزیع کی دو t^2 : تابع چگالی کی دو عبارت است از:

$$f(t^2) = c(t^2)^{\frac{\epsilon-2}{2}} \times e^{-\frac{x^2}{2}}$$

این قانون تنها از یک پارامتر ϵ یعنی درجه آزادی تبعیت می‌کند و C یک عدد ثابت وابسته به ϵ است و طوری تعیین می‌شود که سطح زیر منحنی معادل یک گردد.

امید ریاضی و واریانس این توزیع برابر است با:

$$D(t^2) = 2\epsilon$$

$$E(t^2) = \epsilon$$

و



عناوین فصل سوم

- برآورد
- برآورد مناسب
- میزان اریب
- برآورد نقطه ای
- برآورد نااریب

Payam Noor University Ebook

PNUeb

علی محمد جوادی، دستیار آموزشی دانشگاه پیام
نور مرکز گیلانغرب

۱۰۹



عناوین فصل سوم

- تخمین زن
- کاراترین تخمین زن
- برآورد فاصله‌ای
- برآورد نقطه‌ای میانگین
- برآورد فاصله‌ای میانگین



عناوین فصل سوم

- برآورد تفاضل دو میانگین
- برآورد نسبت
- برآورد فاصله‌ای تفاضل دو نسبت
- برآورد واریانس
- برآورد ضریب همبستگی



بر آورد

تعیین تقریبی مقدار پارامتر یا پارامترها توسط نمونه تصادفی به حجم n ، بر آورد کردن یا تخمین زدن آماری نامیده می شود.



برآورد مناسب

برای آنکه برآورد پارامتر از جامعه ، برآورد مناسبی باشد بایست :

اولاً واریانس برآورد کم باشد ،

ثانیاً برآورد نااریب باشد .



میزان اریب

تفاضل بین امید ریاضی برآورد کننده و کمیت مورد برآورد جامعه را میزان اریب گویند. اگر این تفاضل صفر باشد برآورد کننده را ناریب و در غیر اینصورت آن را برآورد کننده اریب می گویند. به عبارتی اگر امید ریاضی یک پارامتر برابر با پارامتر متناظر جامعه باشد آن پارامتر را ناریب می نامند.

مقدار مشخصه جامعه - (برآورد کننده) E = میزان اریب



برآورد نقطه‌ای

برآوردی از یک پارامتر جامعه که بایک عدد مشخص گردد برآورد نقطه‌ای آن پارامتر نامیده می‌شود.

معمولاً پارامتر جامعه را با " و برآورد نقطه‌ای آن را با [^]" نمایش می‌دهند.



برآورد کننده نااریب

برآورد کننده نااریب $\hat{\mu}$ است اگر و فقط اگر داشته باشیم

$$E\left(\hat{\mu}\right) = \mu$$

مثلاً اگر میانگین جامعه \sim و واریانس جامعه μ^2 باشند و از این جامعه

نمونه تصادفی n تایی انتخاب گردد که مقادیر آنها x_1, x_2, \dots, x_n

شود آن گاه $\bar{x} = \frac{\sum x_i}{n}$ برآوردی \sim و $s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$ برآوردی برای μ^2 هستند.



تخمین زن

تخمین زن دستور یا قاعده ای است که نشان می‌دهد چگونه یک تخمین را بر اساس مقادیر به دست آمده در نمونه باید محاسبه کرد. مثلاً در مورد میانگین فرمول آن تخمین زن است.



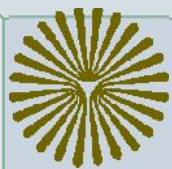
کاراترین تخمین زن

اگر برای یک پارامتر جامعه چند تخمین زن (فرمول) نااریب وجود داشته باشد تخمین زنی که دارای کمترین واریانس باشد را تخمین زن دقیق و یا کاراترین تخمین زن می گویند



برآورد فاصله‌ای

وقتی که برآورد یک مشخص کننده یا پارامتر توسط دو عدد نشان داده شود آن مشخص کننده بین آن دو عدد واقع است. در اینصورت برآورد را فاصله ای می‌گویند. چون دقت و صحت برآورد فاصله ای بیشتر از برآورد نقطه ای است بدین دلیل بر برآورد نقطه ای برتری دارد.



برآورد فاصله ای

اگر به برآورد نقطه‌ای هر پارامتر خطای معیار آن را اضافه و کم کنیم برآورد فاصله ای به دست می‌آید که کرانه بالا را حد بالا (\bar{L}) و کرانه پایین را حد پایین (\underline{L}) فاصله اطمینان می‌نامند. در واقع [^] بین آن دو حد قرار می‌گیرد. به عبارتی :

خطای معیار همان پارامتر \pm برآورد نقطه‌ای پارامتر = برآورد فاصله ای پارامتر

Payam Noor University Ebook

PNUeb

علی محمد جوادی، دستیار آموزشی دانشگاه پیام نور مرکز گیلانغرب

۱۲۰



نکته

اگر حجم نمونه (n) افزایش یابد برآورد ما به مقدار پارامتر نزدیکتر بوده وبالأخره فاصله برآورد کوتاهتر خواهد بود. هرچه فاصله برآورد کوتاهتر گردد دقت برآورد بیشتر خواهد بود .



نکته

فاصله اطمینان بیشتر به ما اطمینان زیادتری خواهد داد که فاصله داده شده شامل پارامتر مجهول خواهد بود.



نکته

در حالت متعارف ترجیح خواهیم داد که فاصله کوتاهتر با درجه اطمینان بیشتر داشته باشیم.



بر آوردن نقطه‌ای میانگین

بر آوردن نقطه‌ای هر میانگین از داده های نمونه طبق فرمول زیر به دست می آید .

$$\bar{x} = \frac{\sum x_i}{n}$$

$$\sim = E(\bar{x})$$

و بدون اریب می باشد.



برآورد فاصله ای میانگین

برآورد فاصله ای میانگین از رابطه زیر به دست می آید $\bar{x} \pm d$: ~
که در آن d را خطای معیار \bar{x} یا خطای نمونه می گویند و مقدار آن
برای نمونه های کوچک برابر است با

$$d = t \sqrt{\frac{SS}{n(n-1)}} \quad \text{و یا} \quad d = t \frac{s}{\sqrt{n}}$$

$$SS_x = \sum x_i^2 - \frac{(\sum x_i)^2}{n} \quad \text{از طرفی}$$

علی محمد جوادی، دستیار آموزشی دانشگاه پیام
نور مرکز گیلانغرب



برآورد فاصله ای میانگین

در صورتی که حجم نمونه به اندازه کافی بزرگ باشد ($n \geq 30$)
(تقریباً نرمال باشد)

با ۵ درصد خطا مقدار d برابر است با:

$$d = 1.96 \times \frac{s}{\sqrt{n}}$$

و با یک درصد خطا مقدار d برابر با

$$d = 2.58 \times \frac{s}{\sqrt{n}}$$



برآورد فاصله ای میانگین

برای جامعه نرمال فاصله اطمینان برابر است با :

$$\tilde{~} = \bar{x} \pm z \frac{u}{\sqrt{n}}$$



برآورد فاصله ای میانگین

در صورتی که U مجهول باشد به عبارتی توزیع X نرمال نباشد آنگاه فاصله اطمینان برابر است با:

$$\sim = \bar{x} \pm t \frac{s}{\sqrt{n}} = \bar{x} \pm t \left(\bar{x} \text{ معیار خطای } \bar{x} \right)$$



برآورد تفاضل دو میانگین

در صورتیکه برآورد تفاضل دو میانگین واقعی $(\bar{\mu}_1 - \bar{\mu}_2)$ مورد نظر باشد آن گاه خطای معیار متغیر تصادفی $(\bar{x}_1 - \bar{x}_2)$ عبارت است از

$$d = t \cdot \sqrt{\frac{SS_1 + SS_2}{n_1 + n_2 - 2} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$



نکته

با افزایش حجم نمونه (n) و همچنین با کاهش انحراف معیار (u و یا S) خطای نمونه‌گیری (d) کاهش می‌یابد.



تبصره

اگر اطمینان داشته باشیم که واریانس واقعی دو جامعه مورد مطالعه یکسان نباشد آن گاه خطای معیار برابر است با :

$$s_d = \sqrt{\frac{SS_1}{n_1(n_1 - 1)} + \frac{SS_2}{n_2(n_2 - 1)}}$$



برآورد نسبت

نسبت واحدهای جامعه که ویژگی مورد نظر را دارا هستند با علامت f نشان خواهیم داد .

و p مساوی است با تعداد افراد نمونه که یک ویژگی به خصوص را دارا هستند تقسیم بر تعداد کل افراد نمونه .

P برآوردی است نااریب از f (نسبت واقعی در جامعه) یعنی:

$$E(p) = f$$



برآورد نسبت

واریانس حقیقی متغیر تصادفی p عبارت است از

$$u^2 = \frac{f(1-f)}{n}$$



برآورد نسبت

s_p^2 برآوردی از واریانسی است که در جامعه وجود دارد که آنرا به شکل زیر نشان می‌دهند:

$$s_p^2 = \frac{p(1-p)}{n}$$

$$s_p^2 = \frac{pq}{n}$$



برآورد نسبت

برآورد فاصله‌ای یا فاصله اطمینان برای f عبارتست از :

$$p - d < f < p + d$$

که در فاصله اطمینان ۹۵ درصد d برابر است با :

$$d = z \cdot s_p = 1/96 \sqrt{\frac{pq}{n}}$$



برآورد فاصله ای تفاضل دو نسبت

اگر نسبت صفت A را در نمونه ای از جامعه ۱، p_1 و در نمونه ای از جامعه ۲، p_2 بنامیم. آنگاه انحراف معیار متغیر تصادفی $(p_1 - p_2)$ برابر است با:

$$s_{(p_1 - p_2)} = \sqrt{\frac{p_1(1 - p_1)}{n_1} + \frac{p_2(1 - p_2)}{n_2}}$$

معمولاً برای سادگی $s_{p_1 - p_2}$ را با s_p نشان می دهند.



برآورد فاصله ای تفاضل دو نسبت

فاصله اطمینان برای تفاضل نسبت های دو جامعه $(f_1 - f_2)$ با ۹۵ درصد اطمینان برابر است با :

$$(f_1 - f_2) : (p_1 - p_2) \pm 1/96 \times s_p$$

که در آن f_1 و f_2 نسبت واقعی در جامعه های اول و دوم هستند، که مجهول می باشند.



برآورد واریانس

برآورد نقطه ای واریانس از داده های نمونه طبق فرمول زیر به

$$SS = \sum x_i^2 - \frac{(\sum x_i)^2}{n}$$

دست می آید. $s^2 = \frac{SS}{n-1}$ که در آن

در صورتیکه صفت کیفی باشد واریانس نسبت برابر است با :

$$s_p^2 = \frac{p(1-p)}{n} = \frac{pq}{n}$$

Payam Noor University Ebook

علی محمد جوادی، دستیار آموزشی دانشگاه پیام
نور مرکز گیلانغرب



برآورد واریانس

واریانس نقطه‌ای مجموع ویا تفاضل دونسبت نیز برابر است با :

$$s_p^2 = \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}$$



برآورد واریانس

برآورد فاصله ای واریانس واقعی جامعه u^2 از نامساوی زیر به دست می آید.

$$\frac{ns^2}{t_{1-\frac{r}{2}}^2} < u^2 < \frac{ns^2}{t_{\frac{r}{2}}^2}$$

که در آن n حجم نمونه و s^2 برآورد نقطه ای واریانس است و $t_{\frac{r}{2}}$ ، $t_{1-\frac{r}{2}}$ با درجه آزادی $n-1$ از جدول توزیع کی دو بدست می آید و مقدار آن به درصد اطمینان مربوط است r نیز میزان درصد خطاست.

Payam Noor University Ebook

PNUeb

علی محمد جوادی، دستیار آموزشی دانشگاه پیام نور مرکز گیلانغرب



برآورد ضریب همبستگی ...

در عمل به علت محدود بودن مشاهدات محاسبه ضریب همبستگی واقعی ... بین دو متغیر X, Y امکانپذیر نیست به این دلیل باید آنها را از روی نمونه ها برآورد کرد. برآوردگر ... را با r نمایش می دهند.



برآورد ضریب همبستگی ...

چند فرمول از ضریب همبستگی پیرسون :

$$r = \sqrt{aa'}$$

$$r = \frac{\text{cov}(X, Y)}{u_X \cdot u_Y}$$

$$r = \frac{sp}{\sqrt{SS_X \cdot SS_Y}}$$

$$r = a \frac{u_x}{u_y}$$



برآورد ضریب همبستگی ...

گرچه r یک برآورد کننده اریب از ρ می باشد ولی در عمل همیشه به عنوان برآورد کننده ρ انتخاب می شود. هر قدر حجم نمونه n کاهش یابد اریبی بیشتر می شود ولی وقتی n به اندازه نامتناهی بزرگ شود اریب از بین می رود.



عناوین فصل چهارم

فرض آماری

انواع خطادر استنباط

داده های پارامتری و ناپارامتری

آزمون توزیع نرمال

آزمون t استودنت

آزمون یک دامنه و دو دامنه

Payam Noor University Ebook

PNUeb

علی محمد جوادی، دستیار آموزشی دانشگاه پیام
نور مرکز گیلانغرب

۱۴۴



عناوین فصل چهارم

📖 آزمون F یا تجزیه و تحلیل واریانس

📖 تفسیر آزمون F

📖 گروه بندی جامعه های مورد مطالعه

📖 رابطه آزمون F با آزمون t

📖 کاربرد توزیع t^2

📖 محاسبه فراوانیهای مورد انتظار (یا تئوریک)



عناوین فصل چهارم

📖 درجه آزادی

📖 t^2 قضاوت آزمون

📖 آزمون t^2 برای جدول دو بعدی (توافقی)

📖 تصحیح یتس

📖 ادغام سطرها و ستون ها



فرض آماری

هر فرض در مورد پارامترهای نامعلوم (میانگین، واریانس، و...) یک جامعه آماری را فرض آماری می‌گوییم.

فرض آماری قاعده یا دستوری است که بر اساس نمونه انتخاب شده به دست آمده و بر مبنای آن فرضیه مورد نظر را قبول یا رد می‌نماییم

Payam Noor University Ebook

PNUeb

علی محمد جوادی، دستیار آموزشی دانشگاه پیام
نور مرکز گیلانغرب

۱۴۷



فرض صفر

منظور از فرض صفر این است که تفاضل دو پارامتر مورد مطالعه قابل ملاحظه نیست به عبارت دیگر اختلاف چندانی بین پارامتر بدست آمده از نمونه و پارامتر مورد نظر ما ، مشاهده نمی شود و می توان گفت این دو پارامتر تقریباً برابرند.



رض H_1

منظورمان از فرض H_1 این است که دو پارامتر مورد مطالعه یکسان نبوده و دارای اختلاف معنی داری باشند به عبارت دیگر تفاوت آنچه مشاهده شده با نتایج مورد انتظار، زیاد می باشد.

انواع خطا در استنباط آماري

الف-خطای نوع اول (α): اگر به اشتباه، فرض H_0 را (که باید قبول شود) رد کنیم، مرتکب «خطای نوع اول» شده ایم.

انواع خطا در استنباط آمار

ب- خطای نوع دوم (β): اگر به اشتباه، فرض نادرست H_0 را (که باید رد شود) قبول کنیم «خطای نوع دوم» روی داده است.



نکته

در تحقیقات آماری ، تعیین میزان خطای نوع اول α برخطای نوع دوم S مقدم است و این سطح احتمال را «سطح اعتماد» یا «سطح معنی دار بودن» می گویند.



داده های پارامتری

هر متغیری که بتوانیم مقدار آن را اندازه گیری کنیم ، داده های آن متغیر را «داده های پارامتری می نامیم. نرمال بودن اینگونه داده ها تا حدی الزامی است.

آزمونهای پارامتری : آزمونهای U, t و F را آزمونهای پارامتری می نامیم.



موارد کاربرد آزمونهای پارامتری

- هریک از نمونه‌ها مستقل بوده و وابسته به هم نباشند .
- واریانس نمونه‌ها برابر یا تقریباً برابر باشند
- اندازه‌گیری آنها با استفاده از مقیاس فاصله‌ای یا نسبی انجام شود. داده‌های اسمی (شمارش افراد) و ترتیبی (رتبه بندی) برای آزمونهای پارامتری مناسب نیستند.



داده‌های ناپارامتری

داده‌های ناپارامتری یا قابل شمارش بوده و یا رتبه بندی می‌شوند. در اینجا متغیر به صورت کیفی است در نتیجه طبقه بندی شده و یا بر حسب فراوانی ارائه می‌شوند بنابراین به پیش فرض نرمال بودن توزیع جامعه‌ها استوار نیستند.

آزمونهای ناپارامتری: آزمون t^2 و M.W (من-وایت نی)

موارد کاربرد آزمون‌های ناپارامتری

- نرمال بودن جامعه ای که نمونه از آن انتخاب می‌شود، معلوم نباشد.
- متغیر به صورت کیفی باشد (اعم از اینکه رتبه‌ای یا غیر رتبه‌ای باشد).



مطالبی از منحنی نرمال

📖 معادله منحنی نرمال که دارای متغیر استاندارد شده است به

$$y = \frac{1}{\sqrt{2f}} e^{-\frac{u^2}{2}}$$

شرح زیر می باشد:

که در آن $U = \frac{x_i - \bar{x}}{u}$ می باشد و شرایط زیر برقرار است:

$$f = 3.14 \quad \text{و} \quad e = 2.72 \quad \text{و} \quad -\infty < U < +\infty$$



شکل منحنی نرمال

WWW*PNUeB*COM

Payam Noor University Ebook

PNUeB

علی محمد جوادی، دستیار آموزشی دانشگاه پیام
نور مرکز گیلانغرب

۱۵۸



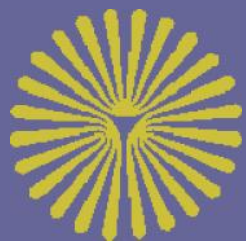
مطالبی از منحنی نرمال

با توجه به اینکه میانگین این منحنی صفر و انحراف معیار آن یک است، فرم ریاضی منحنی نرمال استاندارد شده را به صورت $N(0,1)$ می نویسند.

این منحنی را «منحنی گاوس»، «منحنی خطاها» و به دلیل شباهت آن به ناقوس آن را «منحنی زنگی لاپلاس» نیز می نامند. و چون:

$$\frac{1}{\sqrt{2f}} \int_{-\infty}^{+\infty} e^{-\frac{u^2}{2}} du = 1$$

این منحنی را «منحنی احتمالات» می گویند.



مطالبی از منحنی نرمال

📖 در منحنی نرمال $28/68$ درصد از سطح زیر منحنی بین $U = \pm 1$ ،
 $45/95$ درصد آن بین $U = \pm 2$ و بالاخره $73/99$ درصد آن بین
 $U = \pm 3$ واقع شده است.

📖 به ازای $U = \pm 64/1$ تقریباً 10 درصد، به ازای $U = \pm 96/1$ تقریباً
 5 درصد و به ازای $U = \pm 58/2$ تقریباً یک درصد از کل مساحت
زیر منحنی در دو طرف U قرار می گیرند.

📖 نواحی که در دو طرف هریک از U های ذکر شده قرار می گیرند
ناحیه بحرانی یا ناحیه رد فرض نامیده می شود.



آزمون توزیع نرمال

صفت متغیر جامعه بر طبق قانون توزیع نرمال با میانگین μ و انحراف معیار σ توزیع شده است (که معمولاً این مقادیر مجهول هستند) از این جامعه نمونه‌ای به حجم n انتخاب شده که میانگین آن \bar{x} می‌باشد. می‌خواهیم بدانیم که آیا بین میانگین نمونه \bar{x} و μ تفاوت معنی داری وجود دارد؟



آزمون توزیع نرمال

در واقع می خواهیم یکی از دو فرض زیر را قبول کنیم :

$$H_0 : m = \sim$$

$$H_1 : m \neq \sim$$

به عنوان ملاک آزمون از کمیت U می توان استفاده نمود.

$$U = \frac{|m - \sim|}{\frac{u}{\sqrt{n}}}$$

علی محمد جوادی ، دستیار آموزشی دانشگاه پیام
نور مرکز گیلانغرب



آزمون توزیع نرمال

U بدست آمده از قسمتهای قبل رابا U_c نمایش می دهیم .
📖 اگر $|U_c| < 1.96$ آنگاه فرض U_c مورد قبول واقع می شود
یعنی میانگین حقیقی جامعه با ۹۵ درصد اطمینان قابل قبول
است .

📖 اگر $|U_c| > 1.96$ آنگاه فرض مساوی بودن m و \sim را رد
می کنیم به عبارت دیگر فرض H_1 مورد تأیید قرار می گیرد.



نکات

📖 آزمون توزیع نرمال را زمانی انجام می دهند که u معلوم بوده و یا حجم نمونه از ۳۰ بزرگتر باشد. ($n > 30$)

📖 در تحقیقات اقتصادی - اجتماعی معمولاً با ۹۵ درصد اطمینان (۵ درصد خطا) قضاوت می کنیم.

📖 اگر U_c بین ۹۶/۱ و ۵۸/۲ قرار گیرد ($۹۶/۱ < U_c < ۵۸/۲$) بهتر است از اتخاذ تصمیم خودداری نمود و برای تصمیم قطعی حجم نمونه را افزایش داد و یا فقط با ۵ درصد خطا قضاوت کنیم.



آزمون t استودنت

آزمون t نیز آزمونی برای بررسی وجود تفاوت میان \bar{x} (میانگین جامعه نمونه) و \sim (عدد فرضی یا معین) و یا بین \sim_1 و \sim_2 (میانگین جامعه‌های نمونه) بر حسب واحد خطای معیار S_x می‌باشد. اگر حجم نمونه از ۳۰ کمتر باشد از توزیع t استفاده می‌کنیم.

$$t = \frac{|\bar{x} - \sim|}{\frac{s}{\sqrt{n}}}$$

و در آن $s = \sqrt{\frac{SS}{n-1}}$ است.

علی محمد جوادی، دستیار آموزشی دانشگاه پیام نور مرکز گیلانغرب



شیوه استفاده جدول t

معادله منحنی t فقط تابع حجم نمونه است و به همین دلیل از جدول t که شامل سطوح مختلف اعتماد و درجات آزادی است استفاده می‌کنیم. شیوه استفاده از جدول t شبیه جدول ضرب است. برای تعیین مقدار t، کافی است که درجه آزادی $df=n-1$ و سطح معنی دار بودن r مشخص باشد.



شیوه استفاده جدول t

تفسیر آن مانند تفسیر آزمون نرمال است. به این ترتیب که اگر t محاسبه شده از t جدول کوچکتر باشد در آن صورت فرض H_0 را می‌پذیریم و در غیر اینصورت فرض H_1 مورد تأیید ماست.



حالات مختلف آزمون t

📖 مقایسه میانگین نمونه (m یا \bar{x}) یک عدد فرضی \sim

📖 مقایسه میانگین دو جامعه (\sim_1 و \sim_2)

📖 مقایسه نسبتی که از نمونه بدست آمده (p) و یک

نسبت فرضی (f)

📖 مقایسه دو نسبت از دو جامعه (f_1 و f_2)



آزمون یک دامنه و دو دامنه

اگر هدف اصلی آزمونی تعیین اختلاف دو مقدار یا دو نسبت باشد و به جهت تغییرات آن یعنی مثبت و یا منفی بودن، افزایش یا کاهش داشتن، کمتر و یا بیشتر بودن، بزرگتر و یا کوچکتر بودن و جملاتی نظیر آنها توجه نکنیم، آزمون ما دو دامنه خواهد بود. در غیر اینصورت اگر جهت تغییرات مورد نظر باشد آزمون یک دامنه است.



آزمون یک دامنه و دو دامنه

وقتی از صورت مسأله به یک دامنه بودن آزمون پی بردیم آنگاه باید به جای ستون ۵ درصد خطا در جدول t ، به ستون ۱۰ درصد خطا مراجعه نماییم. اما در مورد تفسیر، باز هم با ۵ درصد خطا قضاوت می‌کنیم زیرا این جدول برای آزمون دو دامنه است.



حالت اول آزمون t

این حالت مقایسه میانگین جامعه با یک مقدار فرضی است .

$$t = \frac{d}{s_d}$$

فرمول حالت اول t استودنت را چنین نوشت :

$$s_d = \sqrt{\frac{SS}{n(n-1)}} \quad \text{و یا} \quad s_d = \frac{s}{\sqrt{n}} \quad \text{و} \quad d = |\bar{x} - \mu|$$

و درجه آزادی برابر با $d.f = n - 1$ می باشد.



حالت دوم آزمون t

این حالت برای بیان تفاوت یا عدم تفاوت بین میانگین‌های دو جامعه است .

اگر واریانس حقیقی دو جامعه یکسان باشد می‌توان انحراف معیار متغیر تصادفی d را از رابطه زیر تعیین کرد :

$$s_d = \sqrt{\frac{SS_{x1} + SS_{x2}}{n_1 + n_2 - 2} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

و درجه آزادی آن $d.f = n_1 + n_2 - 2$ است.



حالت دوم آزمون t

و در صورتی که واریانس اصلی دو جامعه یکسان نباشد :

$$s_d = \sqrt{\frac{SS_{x1}}{n_1(n_1 - 1)} + \frac{SS_{x2}}{n_2(n_2 - 2)}}$$

$d.f_1 = n_1 - 1$ و $d.f_2 = n_2 - 1$ ؛ در نتیجه از جدول دو t به دست می آید.



حالت دوم آزمون t

اگر t_c از هر دو t جدول بزرگتر یا کوچکتر باشد مانند قبل قضاوت می‌کنیم ولی اگر t محاسبه شده بین دو t جدول قرار گیرد t محاسبه شده با t متوسط (میانگین وزنی) بر طبق فرمول زیر محاسبه می‌شود و مانند قبل قضاوت می‌شود. اگر H_1 مورد قبول واقع شد باید دید که کدام میانگین بزرگتر است آنگاه آنرا تعبیر و تفسیر نمود.

$$t_M = \frac{(n_1 - 1)t_1 + (n_2 - 1)t_2}{n_1 + n_2 - 2}$$

Payam Noor University Ebook

PNUEB

علی محمد جوادی، دستیار آموزشی دانشگاه پیام نور مرکز گیلانغرب



حالت سوم آزمون t

این حالت یک عدد فرضی را با یک نسبت که از نمونه های تصادفی به دست آمده مقایسه می کند.

در اینجا ملاک آزمون عبارت است از $t = \frac{d}{s_p}$ که در آن

$$s_p = \sqrt{\frac{f(1-f)}{n}} \text{ و } d = |p - f|$$

همچنین درجه آزادی

برابریست با $d.f = \infty$



حالت سوم آزمون t

📖 چون f از P به واقعیت نزدیکتر است لذا در محاسبه انحراف معیار به جای P نسبت f را قرار می دهند



حالت چهارم آزمون t

مقایسه نسبت‌های بدست آمده از نمونه‌های تصادفی مستقل از دو جامعه مختلف را انجام می‌دهد.

می‌دانیم که: $t = \frac{d}{s_p}$ از طرفی $d = |p_1 - p_2|$ و

$$s_p = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}} \quad \text{و} \quad d.f = \infty$$

آزمون F یا تجزیه و تحلیل واریانس

- 📖 آزمون F دو یا چند میانگین را مقایسه می کند .
- 📖 قانون توزیع کمیت تصادفی F از دو پارامتر درجه آزادی $d.f_1$ و $d.f_2$ تبعیت می کند .

آزمون F یا تجزیه و تحلیل واریانس

اساس آزمون F براین اصل استوار است که در یک آزمایش واریانس کل جامعه ها به واریانس بین گروهها و واریانس درون گروهها (داخل گروهها) تقسیم می شود.

واریانس درون گروهها + واریانس بین گروهها = واریانس کل

آزمون F یا تجزیه و تحلیل واریانس

واریانس درون گروهها: پراکندگی مقادیر متغیرها را درون هر یک از k گروه نشان می دهد و آن را واریانس خطای نامند و SS مربوط به آن را با SS_e نشان می دهند.

واریانس بین گروهها: اندازه اختلاف بین میانگین های k نمونه را نشان می دهد و SS مربوط به آن را نیز با SS_r نشان می دهند

آزمون F یا تجزیه و تحلیل واریانس

در آزمون F:

📖 SS_T به دو بخش SS_e و SS_r تقسیم می شود:

$$SS_T = SS_e + SS_r$$

📖 و درجه آزادی کل df_T نیز به دو بخش df_e و df_r تقسیم

می شود:

$$df_T = df_e + df_r$$

آزمون F یا تجزیه و تحلیل واریانس

$$SS_T = \sum_{i=1}^N x_i^2 - CF_T$$

که در آن $CF_T = \frac{(\sum x_i)^2}{N}$ و N مجموع حجم نمونه‌هاست.

$$SS_r = \sum_{i=1}^k cf_i - CF_T$$

که در آن $\sum CF_i$ مجموع جمله دوم SS هر یک از گروه‌های مورد مطالعه است.

آزمون F یا تجزیه و تحلیل واریانس

چون بین SS_T و SS_e و SS_r رابطه $SS_T = SS_e + SS_r$ برقرار است کفایت SS_r و SS_T را داشته باشیم .

آزمون F یا تجزیه و تحلیل واریانس

📖 درجه آزادی کل عبات است از: $d.f_T = N - 1$

📖 درجه آزادی بین گروهها (n_1) عبارت است از:

$$d.f_r = k - 1$$

📖 درجه آزادی درون گروهی (n_2) برای k گروه عبارتست از:

$$df_e = \sum_{i=1}^k n_i - k = N - k$$

آزمون F یا تجزیه و تحلیل واریانس

متوسط مربعات بین گروهی و درون گروهی :

$$MS_e = \frac{SS_e}{df_e} \quad \text{و} \quad MS_r = \frac{SS_r}{df_r}$$

آزمون F یا تجزیه و تحلیل واریانس

برای اینکه درباره تفاوت حقیقی بین جامعه های جزء قضاوت نماییم MS_r را به MS_e تقسیم می کنیم و نسبت حاصل را با حرف F نمایش می دهیم.

$$F = \frac{MS_r}{MS_e}$$

البته معمولاً MS_r بزرگتر از MS_e می باشد.

آزمون F یا تجزیه و تحلیل واریانس

جدول

WWW*PNUeB*COM

Payam Noor University Ebook

PNUeB

علی محمد جوادی، دستیار آموزشی دانشگاه پیام
نور مرکز گیلانغرب

۱۸۷



تفسیر آزمون F

تفسیر آزمون F مانند سایر آزمونها است.

اگر فرض H_1 مورد قبول واقع شود بین گروهها اختلاف واقعی وجود دارد.

یعنی میان واریانس بین گروهها و درون گروهها اختلاف چشمگیری وجود دارد.

نکته



📖 در آزمون F بحث یک دامنه و دو دامنه بودن
اصلاً مطرح نیست .

Payam Noor University Ebook

PNUeb

علي محمد جوادى ،دستيار آموزشى دانشگاه پیام
نور مرکز گیلانغرب

۱۸۹

گروه‌بندی جامعه‌های مورد مطالعه

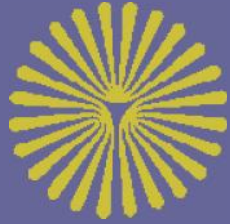
هدف اساسی از گروه‌بندی جامعه‌های مورد مطالعه تقسیم‌بندی جامعه‌ها به دو یا چند گروه به طوری که هر چند جامعه‌ای که در داخل یک گروه قرار می‌گیرند از نظر صفت مورد بررسی یکسان محسوب شوند.

گروه بندی جامعه های مورد مطالعه

برای گروه بندی از کمیتی به نام L.S.D استفاده می کنند :

$$L.S.D = t_{(n_2, \%5)} \times \sqrt{\frac{2MS_e}{n}}$$

که در آن t عددی است که از جدول t استودنت با درجه آزادی n_2 و سطح اعتماد ۵٪ استخراج می شود و MS_e متوسط مربعات اشتباهات و n حجم نمونه است .



نکته

اگر n ها در جامعه های جزء با هم مساوی نباشند می توان به شرط زیاد نبودن اختلاف n ها به جای n از میانگین همساز n_i استفاده کرد .

$$\bar{n}_h = \frac{K}{\sum_{i=1}^m \frac{1}{n_i}}$$

K تعداد گروهها ست.

گروه بندی جامعه های مورد مطالعه

📖 اگر فرض H_1 مورد قبول واقع شد برای گروه بندی جامعه های مورد مطالعه ابتدا میانگین آنها را با توجه به هدفی که داریم بر حسب صعودی یا نزولی مرتب می کنیم. سپس اگر میانگینها بر حسب نزولی مرتب شده باشد، مقدار $L.S.D$ را از بزرگترین میانگین کم می کنیم. $\bar{X}_{Max} - L.S.D$ میانگین هایی که مساوی یا بزرگتر از مقدار فوق باشند در یک گروه قرار می گیرند. پس از کنار گذاشتن میانگین های گروه اول، عملیات را برای میانگین های باقیمانده ادامه می دهیم تا کلیه جامعه ها گروه بندی شوند.

Payam Noor University Ebook

علی محمد جوادی، دستیار آموزشی دانشگاه پیام نور مرکز گیلانغرب

۱۹۳

گروه بندی جامعه های مورد مطالعه

📖 اگر میانگین ها بر حسب صعودی مرتب شوند آن گاه L.S.D را با کوچکترین میانگین جمع می کنیم:

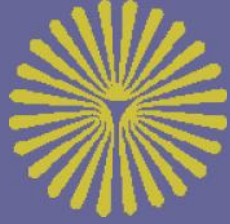
$$\bar{X}_{Min} + L.S.D$$



رابطه آزمون F با آزمون t

آزمون F تعمیم یافته آزمون t است، پس برای مقایسه دو میانگین از هر دو توزیع می توان استفاده کرد. در این صورت در آزمون F درجه آزادی بین گروهها برابر با یک است و بین این دو توزیع رابطه زیر برقرار است :

$$F=t^2$$



کاربرد توزیع t^2

هدف از اجرای آزمون کی دو این است که آیا بین فراوانیهای مشاهده شده (n_i) و فراوانیهای مورد انتظار (n_{ith}) ، تفاوت منظم و معنی داری وجود دارد؟ منظور از مقایسه کردن فراوانیها، تشخیص وابستگی یا عدم وابستگی دو متغیر مورد مطالعه است.



کاربرد توزیع t^2

از آنجا که در آزمون t^2 فراوانیها مورد استفاده قرار می گیرند، بنابراین این آزمون برای صفات کیفی بکار برده می شود که داده ها به صورت شمرده هستند یعنی مقیاس اندازه گیری داده ها اسمی یا ترتیبی است .



ملاک آماری آزمون کی دو

ملاک آزمون t^2 عبارتست از :

$$t^2 = \sum \frac{(n_i - n_{ith})^2}{n_{ith}}$$

که در آن n_i فراوانیهای مشاهده شده و n_{ith} فراوانیهای مورد انتظار یا تئوریک می باشد .

هرچه مشابهت بین n_i و n_{ith} بیشتر باشد مقدار t^2 بدست آمده کوچکتر خواهد بود.



محاسبه فراوانی‌های مورد انتظار

📖 در جداول یک بعدی استفاده از امید ریاضی $E(x) = np$.

📖 استفاده از اطلاعات قبلی که می‌تواند به صورت درصد یا نسبت در اختیار محقق قرار گیرد.

📖 مقایسه توزیع فراوانی‌های مشاهده شده با توزیع فراوانی نظری مانند توزیع نرمال .



درجه آزادی

درجه آزادی در آزمون کی دو به تعداد صفات مورد مطالعه و یا به عبارتی تعداد مقوله‌ها مربوط است .

📖 در جدول یک بعدی اگر K صفت داشته باشیم درجه آزادی برابر است با

$$d.f = K - 1$$

📖 در جدول دو بعدی با K سطرو L ستون درجه آزادی برابر است با

$$d.f = (K - 1) (L - 1)$$



قضاوت آزمون

اگر t^2 محاسبه شده (t_c^2) بزرگتر یا مساوی t^2 جدول

باشد $t_c^2 \geq t_t^2$ فرض (H_0) یا فرض مستقل بودن متغیرها رد می شود.



قضاوت آزمون

📖 در صورتیکه t_c^2 کوچکتر از t^2 جدول باشد $t_c^2 < t_t^2$ آنگاه فرض صفر تأیید می‌شود و نتیجه می‌گیریم که متغیرها مستقل هستند و اختلافات مشاهده شده ناشی از شانس یا خطای نمونه‌گیری است.

آزمون t^2 برای جداول دو بعدی (توافقی)

فراوانیهای مشاهده شده همان ارقام متن جدول ولی فراوانیهای مورد انتظار از فرمول تقسیم به نسبت زیر محاسبه می شود .

$$n_{ith} = \frac{n_{i0} \times n_{0j}}{n}$$

n_{i0} فراوانیهای توزیع حاشیه ای X های یعنی مجموع فراوانیهای سطر i است

n_{0j} فراوانی های توزیع حاشیه ای Y هاست یعنی مجموع فراوانیهای ستون j .



نکته

فرمول t^2 وقتی مورد استفاده است که :

- 📖 حجم نمونه حداقل ۵۰ باشد .
- 📖 فراوانی مورد انتظار برای هر خانه کمتر از ۵ نباشد .



تصحیح یتس

اگر در جدول چهار خانه ای فراوانی های مورد انتظار کمتر از ۵ باشد از تصحیح یتس استفاده می کنیم :

$$t^2 = \sum \frac{[n_i - n_{ith} - 0.5]^2}{n_{ith}}$$

از تصحیح یتس وقتی استفاده می شود که درجه آزادی یک باشد .



ادغام سطرها و ستونها

اگر حجم نمونه کمتر از ۵۰ باشد آنگاه ممکن است فراوانی مورد انتظار کمتر از ۵ باشد. در این حالت اگر جدول داده های یک بعدی باشد صفات متشابه را که فراوانی کمی دارند با هم ادغام می کنیم.

برای جدول دو بعدی که تعداد سطرها و ستون های آن بیش از دو است اگر فراوانی بیش از ۲۰ درصد خانه ها کمتر از ۵ باشد فراوانی سطرها و ستونهای همجوار را در هم ادغام می کنیم.