دِموی جزوه آمار و احتمالات

دانشگاه صنعتی امیر کبیر

دکتر عرب زاده

مثال: پرتاب یک سکه، روشن کردن کلید برق.

متغیر برنولی؟

متغیر تصادفی X که نشان‌دهندهٔ تعداد موفقیت‌ها در آزمایش برنولی باشد را متغیر برنولی گویند.

$X: x=0,1$

موفقیت شکست

$p(S)=p(X=1)=p \quad (0<p<1)$

$p(f)=p(X=0)=1-p=q$

توزیع برنولی

توزیع متغیر برنولی را توزیع برنولی نامیده و با علامت خاص $b(x;p)$ نشان می‌دهند.

x	0	1
$f_X(x)=b(x;p)$	q	p

$f_X(x)=b(x;p)=p^x q^{1-x} \quad x=0,1$

قضیه: میانگین و واریانس توزیع برنولی $b(x;p)$ برابر است با:

$\mu=p \qquad \sigma^2=pq$

اثبات:

الف: با کمک تعریف:

$\mu=E(X)=\sum_{x=0}^{1} x f(x)=0\times q+1\times p=p$

$E(X^2)=\sum_{x=0}^{1} x^2 f(x)=0^2\times q+1^2\times p=p$

$\sigma^2=E(X^2)-(\mu)^2=p-p^2=p(1-p)=pq$

ب: با کمک تابع مولد گشتاور:

$M(t)=E(e^{tX})=\sum_{x=0}^{1} e^{tx} f_X(x)=f(0)+e^t f(1)$

$=q+pe^t$

$M'(t)=$

توزیع دو جمله‌ای (binomial distribution)

آزمایش دو جمله‌ای:

اگر یک آزمایش برنولی را n مرتبه و مستقل تکرار کنیم، این آزمایش را دو جمله‌ای می‌گویند.

مثال: پرتاب یک سکه (سالم یا ناسالم) n مرتبه، روشن کردن n لامپ (لامپ)، خارج کردن n مهره با جایگذاری از ظرفی حاوی a مهره سفید و b مهره سیاه.

متغیر دو جمله‌ای:

متغیر تصادفی X که نشان دهندهٔ تعداد موفقیت در آزمایش دو جمله‌ای می‌باشد را متغیر دو جمله‌ای گویند.

$$X: x = 0,1,2,\dots,n$$

تابع یا توزیع دو جمله‌ای:

توزیع متغیر دو جمله‌ای را توزیع دو جمله‌ای گویند و با علامت $f_X(x)=b(x;n,p)$ نشان می‌دهند

$$f_X(x) = \binom{n}{x} p^x q^{n-x}$$

$$n=1 \Rightarrow f_X(x) = b(x;1,p) = b(x;p)$$

تمرین: تحقیق کنید این تابع یک تابع احتمال است.

$$f_X(x) = b(x;n,p) = \binom{n}{x} p^x q^{n-x}$$

$$\sum_{x=0}^{n} f_X(x) \overset{?}{=} 1 \iff \sum_{x=0}^{n} b(x;n,p) = \sum_{x=0}^{n} \binom{n}{x} p^x q^{n-x} \overset{?}{=} 1$$

$$(a+b)^n = \sum_{r=0}^{n} \binom{n}{r} a^r b^{n-r} \qquad \binom{n}{r} = \frac{n!}{r!(n-r)!}$$

$$\begin{cases} r=x \\ a=p \\ b=q \end{cases} \Rightarrow (p+q)^n = \sum_{x=0}^{n} \binom{n}{x} p^x q^{n-x}$$

قضیه: میانگین و واریانس توزیع دو جمله‌ای $b(x;n,p)$ برابر است با:

$$\mu = np \qquad \sigma^2 = npq$$

اثبات: به کمک تابع مولد گشتاور:

$$M(t) = E(e^{tx}) = \sum_{x=0}^{n} e^{tx} f_X(x) = \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x q^{n-x}$$

$$= \sum_{x=0}^{n} \binom{n}{x} (pe^t)^x q^{n-x}$$ متناظر با بسط دو جمله‌ای

Subject:

Year: Month: Day:

چند خواص از واریانس و کوواریانس

چند قضیه زیر خواص واریانس را نشان میدهد. این خواص به کمک فرمول های ۱ و ۲ به سادگی اثبات میشوند

قضیه ۱:

اگر a و b دو مقدار ثابت باشند، آنگاه نتایج زیر برقرار است:

1) $\sigma^2_{x+b} = \sigma^2_x = \sigma^2$

2) $\sigma^2_{ax} = a^2\sigma^2_x$

3) $\sigma^2_{ax\pm by} = a^2\sigma^2_x + b^2\sigma^2_y \pm 2ab\,\sigma_{xy}$

4) $Cov(ax+b, cy+d) = ac\,Cov(x, y)$

اثبات ۱:

$\sigma^2_{x+b} = E(x+b-\mu_{x+b})^2$

در فرمول ۲ به جای $g(x)$ قرار میدهیم: $x+b$

$\mu_{x+b} = E(x+b) = \mu_x + b$

SAMAND

Subject:

Year: Month: Day:

$\sigma^2_{x+b} = E(x+b-\mu_x-b)^2 = E(X-\mu_x)^2 = \sigma^2_x$

نتیجه: اگر به تمام مقادیر یک متغیر عددی را اضافه کنیم، آن عدد به میانگین اضافه میشود و واریانس تغییر نمیکند.

اثبات ۲: یعنی اگر همه مقادیر یک متغیر را در عددی (مثل a) ضرب کنیم، میانگین a برابر و واریانس a^2

برابر خواهد شد. (اثبات به عهده شما)

$\mu_{ax\pm by} = E(ax \pm by) = a\mu \pm b\mu_y$

$\sigma^2_{ax\pm by} = E(ax \pm by) = a^2\sigma^2_x + b^2\sigma^2_y \pm ab\,\sigma_{xy}$

اگر x و y مستقل باشند:

$\sigma^2_{ax\pm by} = a^2\sigma^2_x + b^2\sigma^2_y$

SAMAND

Subject:
Year: Month: Day:

امید ریاضی شرطی (مشروط)

امید ریاضی متغیر Y به شرط آنکه $X=x$ اتفاق افتاده باشد را به صورت زیر تعریف می‌کنیم.

$$\mu^* = E(Y|X) = \begin{cases} \sum_y y\, f_{Y|X}(y|x) \\ \int y\, f_{Y|X}(y|x)\, dy \end{cases}$$

به همین ترتیب واریانس مشروط به صورت زیر تعریف می‌گردد:

واریانس مشروط متغیر تصادفی Y به شرط آنکه $X=x$ اتفاق افتاده باشد برابر است با

$$V(Y|X) = E\left((Y-\mu^*)^2 \mid X=x\right) = \begin{cases} \sum_y (y-\mu^*)^2 f_{Y|X}(y|x) \\ \int (y-\mu^*)^2 f_{Y|X}(y|x)\, dy \end{cases}$$

مثال/ اگر X و Y دو متغیر تصادفی مستقل با میانگین‌های به ترتیب $\mu_X=2$ و $\mu_Y=3$ و واریانس‌های $\sigma_X^2=1$ و $\sigma_Y^2=2$ باشند، میانگین و واریانس متغیر تصادفی Z به صورت $Z=3X+2Y-7$ را محاسبه کنید.

SAMAND

Subject:
Year: Month: Day:

$$\mu_Z = E(Z) = E(3X+2Y-7)$$
$$= 3E(X)+2E(Y)-7$$
$$= 3\mu_X + 2\mu_Y - 7$$
$$= 3(2)+2(3)-7 = 5 \Rightarrow \mu_Z = 5$$

$$\sigma_Z^2 = \sigma^2_{\underbrace{3X+2Y-7}_{W\ \text{متغیر}}} = \sigma^2_{3X+2Y} = 9\sigma_X^2 + 4\sigma_Y^2 = 9(1)+4(2) = 17$$

مثال/ اگر X و Y دو متغیر تصادفی با تابع احتمال توأم زیر باشند،

$$f_{X,Y}(x,y) = \begin{cases} 8xy & 0\le x\le 1 ,\ 0\le y\le x \\ 0 & \text{در سایر جاها} \end{cases}$$

اولاً بررسی کنید که f یک تابع احتمال توأم است. $\left(\int_0^1\int_0^x 8xy\, dy\, dx\right)$

ثانیاً مطلوبست $\mu^* = E(Y|X)$ و $V(Y|X)$

$$f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)\,dy = \int_0^x 8xy\,dy = \begin{cases} 4x^2 & 0\le x\le 1 \\ 0 & \text{در سایر جاها} \end{cases}$$

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{8xy}{4x^2} = \frac{2y}{x} \qquad 0\le x\le 1$$

$$\mu^* = \int_0^x y \times \frac{2y}{x}\,dy = \frac{2x^2}{3} \qquad 0< x\le 1$$

SAMAND

Subject:
Year: Month: Day:

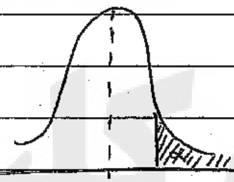
بصورت زیر باشد:

$$\begin{cases} H_0: \theta = \theta_0 \\ H_1: \theta > \theta_1 \end{cases} \quad \text{یا} \quad \begin{cases} H_0: \theta = \theta_0 \\ H_1: \theta < \theta_0 \end{cases}$$

مثال:

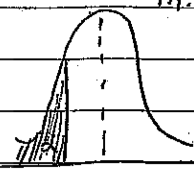
$$\begin{cases} H_0: \mu = 170 \\ H_1: \mu > 170 \end{cases} \quad \text{یا} \quad \begin{cases} H_0: \mu = 170 \\ H_1: \mu < 170 \end{cases}$$

از ویژگی یک آزمون یک طرفه آن است که ناحیه بحرانی آن یک طرفه بوده و در یک سمت توزیع قرار می‌گیرد. (اگر آزمون بصورت $\begin{cases} H_0: \theta = \theta_0 \\ H_1: \theta < \theta_0 \end{cases}$ باشد، ناحیه بحرانی سمت چپ توزیع قرار می‌گیرد. و اگر $\begin{cases} H_0: \theta = \theta_0 \\ H_1: \theta > \theta_0 \end{cases}$ باشد، ناحیه بحرانی در سمت راست توزیع قرار می‌گیرد.



$H_0: \mu_0$

$H_1: \mu > \mu_0$



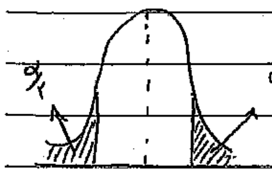
$H_0 = \mu_0$

$H_1: \mu < \mu_0$

* آزمون فرض دوطرفه به آزمون فرضی می‌گویند که فرض مقابل آن دو طرفه باشد، به عبارت دیگر $\begin{cases} H_0: \theta = \theta_0 \\ H_1: \theta \neq \theta_0 \end{cases}$ باشد.

Subject:
Year: Month: Day:

از ویژگی آزمون‌های دوطرفه آن است که ناحیه بحرانی در دو سمت توزیع قرار می‌گیرد.



$H_0: \mu_0$

$H_1: \mu \neq \mu_0$

بطور کلی برای انجام یک آزمون هنگامیکه از توزیع‌های پیوسته استفاده میشود (آزمونهایی که توزیع‌های پیوسته را دارا باشند)، مراحل زیر اجرا میشود.

۱- تشکیل فرض H_0 $\quad H_0: \theta = \theta_0$

۲- تشکیل فرض مقابل $\quad H_1: \theta < \theta_0$ یا $\theta > \theta_0$ یا $\theta \neq \theta_0$

۳- انتخاب سطح تشخیص آزمون (انتخاب مقدار α)

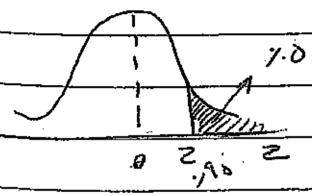
۴- انتخاب آماره مناسب برای θ و تشخیص نوع توزیع آماره یا برآوردگر

۵- تشکیل ناحیه بحرانی به کمک α و توزیع آماره

۶- محاسبات: به کمک نمونه منتخب مقدار آماره مشخص میشود.

Subject:
Year: Month: Day:

۵) توزیع $\bar{x}$ نرمال است چون ($n=100 > 30$) است. یا نرمال استاندارد $Z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$



$Z_{.95}$ = با کمک جدول = ۱/۶۴۵

$Z > 1/645$ ناحیه بحرانی

۶) محاسبات: $n=100$ ، $\bar{x}=1265$ ، $\sigma=300$

$$Z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}} = \frac{1265-1200}{\frac{300}{\sqrt{100}}} = 2/16$$

۷) نتیجه گیری: فرض H_0 رد می‌شود. می‌پذیریم ادعای این تولید کننده درست است.

مثال: میانگین یک نمونه ۱۰۰ تایی از نوعی لامپ دارای طول عمر $\bar{x}_{100}=1570$ و انحراف معیار $\sigma=120$ ساعت است. آزمون فرض را بصورت زیر و در سطح تشخیص $\alpha=0.05$ و $\alpha=0.01$ انجام دهید.

$H_0: \mu=1600$ ← آزمون دو طرفه

$H_1: \mu \neq 1600$

$\alpha=0.05 \longrightarrow \alpha/2=0.025$

SAMAND

Subject:
Year: Month: Day:

۷- نتیجه گیری: اگر مقدار آماره در ناحیه بحرانی قرار گیرد، فرض H_0 رد می‌شود. در غیر این صورت می‌گوییم دلیلی برای رد H_0 پیدا نشد.

مثال: گزارش شده است که میانگین طول عمر نوعی لامپ، $\mu=1200$ ساعت با انحراف معیار $\sigma=300$ ساعت است. کارخانه‌ای مدعی است لامپ جدیدی از این نوع تولید کرده که دارای میانگین طول عمر بیشتر از ۱۲۰۰ ساعت است. برای بررسی این ادعا یک نمونه تصادفی ۱۰۰ تایی از این نوع لامپ جدید انتخاب و مشخص شده است که میانگین طول عمر این نمونه $\bar{x}_{100}=1265$ ساعت است. در مورد این ادعا چه اظهار نظری می‌توان کرد. به عبارت دیگر می‌خواهیم آزمون فرض زیر را انجام داشته باشیم

که $H_0: \mu \le 1200$ ، $H_1: \mu > 1200$ باشد. (سطح تشخیص $\alpha=0.05$ است)

۱) $H_0: \mu=1200$ ← آزمون یک طرفه

۲) $H_1: \mu > 1200$

۳) $\alpha=0.05$ (آزمون یک طرفه)

۴) μ ، آماره مناسب برای میانگین = $\bar{x}$

SAMAND

پایان

جهت دانلود نسخه ی کامل محصول

روی دکمه زیر کلیک نمایید