Subject ---------- Date ----------

استاد کهرلی   آمار:

آزمایش تصادفی: آزمایشی که کلیه نتایج و برآمدهای ممکن آن از قبل معلوم است اما اینکه کدام نتیجه حاصل می‌شود مشخص نیست مانند آزمایش پرتاب یک سکه یا شرایت یا خط اما اینکه کدام یک اتفاق می‌افتد مشخص نیست. مثل مسابقه فوتبال (تساوی، برد، باخت) اما اینکه کدام اتفاق می‌افتد مشخص نیست.

فضای حالت یا فضای نمونه: به کل برآمدهای یک آزمایش تصادفی فضای نمونه می‌گویند. $\{1, 2, \dots, 6\}$ = کلاس

پیشامد: هر زیرمجموعه‌ای از فضای نمونه را یک پیشامد می‌گویند.

مثال: یک تاس پرتاب می‌شود. A برآمد زوج تعریف شده است. پیشامد A را بنویسید.

$A = \{2, 4, 6\}$

احتمال یک پیشامد:

$$P(A) = \frac{n(A)}{n(S)}$$

برای مثال قبل $= \frac{3}{6} = 1/2 = 50\%$

قانون احتمال: 1) $0 \leq P_i \leq 1$ 2) $\sum P_i = 1$

جمع احتمال‌ها

متغیرهای تصادفی: متغیر تصادفی x تابعی است که به هر یک از مقادیر فضای نمونه مقداری عددی نسبت می‌دهد که به دو دسته گسسته و پیوسته تقسیم می‌شود.

شمارش‌پذیر   مرا   مربوط به زمان

توزیع احتمال یک متغیر تصادفی: فهرستی از مقادیر ممکن متغیر تصادفی به همراه احتمال‌های منتسب به آن است.

مثال: در پرتاب یک تاس مطلوب است تعیین جدول توزیع احتمال؟

متغیرهای تصادفی

x_i	1	2	3	4	5	6
$P(x_i)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

احتمال هر کدام

$0 \leq x_i \leq 1$

$\sum x_i = \frac{6}{6} = 1$

elios

Subject .... Date ....

$P[x=2] = 1/6$ $\qquad$ $P[x \geq 4] = p(x=4) + p(x=5) + p(x=6)$

$= \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$

$P[x<2] = p(x=1) = \frac{1}{6}$

مثال: دو تاس پرتاب می‌شود. توزیع احتمال x برابر مجموع ۲ تاس بزرگتر و مساوی با ۱۰ می‌باشد. مطلوب است تعیین توزیع احتمال؟

پیشامد   متغیر

x_i	2	3	4	5	6	7	8	9	10	11	12
$p(x_i)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$6/36$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

$(5,5)\ (4,6)\ (6,4)$ $\quad \frac{n(A)=3}{n(S)=36} \Rightarrow \frac{3}{36}$

$P[x \geq 10] = \frac{3}{36} + \frac{2}{36} + \frac{1}{36} = \frac{1}{6}$

$P[2 \leq x \leq 4] = \frac{1}{36} + \frac{2}{36} + \frac{3}{36} = \frac{1}{6}$

امید ریاضی: $E(x)$

میزان (عدد) مورد انتظار در یک توزیع به نوعی تداعی کننده مرکز ثقل یک توزیع است.

* $E(x) = \sum x_i P(x_i)$ یا $\sum x_i p(x_i)$

مثال: امید ریاضی برآمد پرتاب یک تاس سالم را بدست آورید.

x_i	1	2	3	4	5	6
$p(x_i)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

$$\boxed{\sum x_i p(x_i) = 1\times\frac{1}{6} + 2\times\frac{1}{6} + 3\times\frac{1}{6} + 4\times\frac{1}{6} + 5\times\frac{1}{6} + 6\times\frac{1}{6} = \frac{21}{6} = 3.5}$$

مثال: در مثال قبل (مجموع برآمدهای دو تاس) مطلوب است محاسبه امید ریاضی؟

x_i	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲
$p(x_i)$	۱/۳۶	۲/۳۶	۳/۳۶	۴/۳۶	۵/۳۶	۶/۳۶	۵/۳۶	۴/۳۶	۳/۳۶	۲/۳۶	۱/۳۶
$\sum x_i p(x_i)$	۲/۳۶	۶/۳۶	۱۲/۳۶	۲۰/۳۶	۳۰/۳۶	۴۲/۳۶	۴۰/۳۶	۳۶/۳۶	۳۰/۳۶	۲۲/۳۶	۱۲/۳۶

$$\boxed{E(x) = 7}$$

$$\sum x_i\, p(x_i) = 2 \times \frac{1}{36} + 3 \times \frac{2}{36} + 4 \times \frac{3}{36} \cdots = 7$$

قوانین (خواص) امید ریاضی:

$$E(x \pm \alpha) = E(x) \pm \alpha$$

$$E(ax) = aE(x)$$

$$E(x \pm y) = E(x) \pm E(y)$$

واریانس:

واریانس $\sigma^2 = Var = E(x^2) - E^2(x)$

انحراف معیار $\sigma = \sqrt{var} =$ (معیار پراکندگی است)

برای مثال قبل:

$E^2(x) = (3.5)^2 = 12.25$

$\sigma^2 = var = 25.16 - 12.25 = 2.91$

$\sigma = \sqrt{2.91} = 1.7$

مثال: در مثال فوق مطلوب است $E(3x+5)$

$= 3E(x) + 5 = 3(3.5) + 5 = 15.5$

Subject ن Date

مثال: جدول فروش کالای فروشگاهی به صورت زیر است. متغیر تصادفی x تعداد فروش کالای هفتگی است (طی یک هفته) فروشنده به ازای هر کالا ۵۰ تومان سود می‌کند و هزینه هفتگی آن فروشگاه ۲۰ تومان می‌باشد. امید ریاضی سود فروش هفتگی را به دست آورید.

یعنی در هفته فروش نداریم — یعنی در هفته یکبار — ازای فروش

x	۰	۱	۲	۳	۴	۵
P_{x_i}	۰/۱	۰/۱	۰/۲	۰/۳	۰/۲	۰/۱
$\sum x_i p_i$	۰	۰/۱	۰/۴	۰/۹	۰/۸	۰/۵

$\Rightarrow E(x) = ۲/۷$

x متغیر تصادفی (تعداد فروش هفتگی)

$$E(۵۰x-۲۰) = ۵۰\, E(x) - ۲۰ = ۵۰(۲/۷) - ۲۰ = \boxed{۱۱۵}$$

امید سود

* انتظار فروش هفتگی $E(x)=۲/۷$ می‌باشد.

توزیع توام با هم دو متغیر تصادفی:

هرگاه دو متغیر تصادفی همزمان رخ دهند در جدول توزیع توام به سطر آن یک متغیر و ستون آن متغیر دیگر را می‌باشد استفاده می‌کنیم.

$y \backslash x$	$x_1 \ldots\ldots x_n$
y_1	○ $\rightarrow P(x,y) \;\underline{\vee}\; F(x,y)$
⋮	
y_n	

روش حل سریع توزیع فوق هندسی:

ترکیب: $C_r^n = \binom{n}{r} = \dfrac{n!}{r!\,(n-r)!}$

مثال $= \binom{۵}{۳} = \dfrac{۵!}{۳!\,(۵-۳)!} = \dfrac{۵\times۴\times\cancel{۳!}}{\cancel{۳!}\times۲\times۱} = ۱۰$

(ترکیب سه تایی از ۵ تا)

clips

مثال: در یک کارتن ۱۵ عدد رادیو وجود دارد که ۵ تا آنها معیوب است. ۶ رادیو به تصادف انتخاب می‌کنیم. چقدر احتمال دارد که ۴ رادیو سالم و ۲ رادیو معیوب باشد.

$$\frac{\binom{10}{4}\times\binom{5}{2}}{\binom{15}{6}} = \frac{\frac{10!}{4!\,6!}\times\frac{5!}{2!\,3!}}{\frac{15!}{6!\,9!}}$$

۱۰ ← سالم ، ۴ ← تعداد سالم ، ۵ ← معیوب ، ۲ ← تعداد معیوب ، ۱۵ ← کل ، ۶ ← تعداد انتخاب رادیو

$$\frac{\frac{10\times9\times8\times7\times\cancel{6!}}{4\times3\times2\times\cancel{6!}}\times\frac{5\times4\times\cancel{3!}}{2\times\cancel{3!}}}{\frac{15\times14\times13\times12\times11\times10\times\cancel{9!}}{6\times5\times4\times3\times2\times\cancel{9!}}} = \frac{210\times10}{5005} = 0.41$$

$$\frac{3603600}{720}$$

مثال توزیع توام:

۱۵ ماشین وجود دارد که ۷ عدد از آنها در وضعیت خوب (G)، ۲ عدد دارای نقص در قسمت انتقال قدرت (DT) و شش عدد دیگر دارای نقص در مکانیزم فرمان (DS). فرض کنید ۲ ماشین به صورت تصادفی انتخاب شده‌اند. متغیرهای تصادفی مورد نظر عبارتند از x: ماشین‌هایی که در انتقال قدرت نقص دارند؛ y: ماشین‌هایی که در انتقال فرمان نقص دارند. جدول توزیع احتمال توام آن را بدست آورید.

x \ y	0	1	2
0	A	B	C
1	D	E	X
2	F	X	X

$B = P(0,1)$
یک ماشین سالم و یک ماشین نقص فرمان

$E = P(1,1)$
یک ماشین نقص فرمان و یک ماشین نقص قدرت

Subject .... Date ....

سبز = y   سفید = x

$A = \underset{10}{\left\lfloor \overset{G}{5} \;\; \overset{DT}{2} \;\; \overset{DS}{3} \right\rfloor} \;\; \underset{2}{\left\lfloor \overset{G}{2} \;\; \overset{DT}{0} \;\; \overset{DS}{0} \right\rfloor} = \dfrac{\binom{5}{2}\binom{2}{0}\binom{3}{0}}{\binom{10}{2}} = \dfrac{10}{45} =$

$B = \;〃\; \underset{2}{\left\lfloor \overset{G}{1} \;\; \overset{DT}{0} \;\; \overset{DS}{1} \right\rfloor} = \dfrac{\binom{5}{1}\binom{3}{1}}{45} = \dfrac{15}{45}$

$C = \;〃\; \underset{2}{\left\lfloor \overset{G}{0} \;\; \overset{DT}{0} \;\; \overset{DS}{2} \right\rfloor} = \dfrac{\binom{3}{2}}{45} = \dfrac{3}{45}$

$D = \;〃\; \underset{2}{\left\lfloor 1 \;\; 1 \;\; 0 \right\rfloor} = \dfrac{\binom{5}{1}\binom{2}{1}}{45} = \dfrac{10}{45}$

$E = \;〃\; \underset{2}{\left\lfloor 0 \;\; 1 \;\; 1 \right\rfloor} = \dfrac{\binom{2}{1}\binom{3}{1}}{45} = \dfrac{6}{45}$

$F = \;〃\; \underset{2}{\left\lfloor 0 \;\; 2 \;\; 0 \right\rfloor} = \dfrac{\binom{2}{2}}{45} = \dfrac{1}{45}$

x \ y	0	1	2	توزیع حاشیه‌ای x
0	10/45	15/45	3/45	28/45
1	10/45	6/45	0	16/45
2	1/45	0	0	1/45
توزیع حاشیه‌ای y	21/45	21/45	3/45	45/45 = 1

x	0	1	2
p(x)	28/45	16/45	1/45

y	0	1	2
p(y)	21/45	21/45	3/45

Subject .... Date ....

احتمال اینکه حداکثر یک ماشین خراب در هفته اتفاق بیفتد؟

$P(x,y \le 1)$

$P(x,y=0)+P(x,y=1)$

یعنی احتمال دارد اتصال باشد

$$\frac{10}{45}+\frac{15}{45}+\frac{10}{45}=\frac{35}{45}$$

x y (0,0) (0,1) (1,0)

احتمال اینکه حداقل دو ماشین خراب باشد؟

$P(x,y \ge 2)$

$P(x,y=2)$

$$\frac{3}{45}+\frac{6}{45}+\frac{1}{45}=\frac{10}{45}$$

(0,2) (1,1) (2,0)

احتمال اینکه در بین دو خودروی انتخابی حداکثر یک خودرو دارای نقص در سیستم ترمز باشد؟

$$P(y \le 1)=P(y=0)+P(y=1)=\frac{21}{45}+\frac{21}{45}=\frac{42}{45}$$

(امتحان)

مثال: کارخانه کوچکی روزانه در دو نوبت کار می‌کند. در بررسی الگوی غیبت کارگران این کارخانه متغیرهای تصادفی عبارتند از x: تعداد غایبین نوبت صبح

y: " " " عصر

(سرپرست کارگران بر مبنای تعداد زیادی از مشاهدات متغیر توزیع احتمال را مطابق جدول ذیل بدست آورد)

مطلوب است تعیین هر یک از احتمالای ذیل:

نوبت عصر — تعداد غایبین — x ← نوبت صبح

x \ y	0	1	2	3	
0	0.05	0.05	0.10	0	0.20
1	0.05	0.10	0.25	0.10	0.50
2	0	0.15	0.10	0.05	0.30
	0.10	0.30	0.45	0.15	1.00

۱- توزیع حاشیه‌ای غایبین صبح و عصر:

۲- احتمال اینکه ...

Subject .... Date ....

۳- احتمال اینکه غایبین صبح بیشتر از عصر باشد

۴- " " " " " و عصر با هم برابر باشند

۵- توزیع مجموع غایبین صبح و عصر

۶- احتمال اینکه حداکثر سه غایب داشته باشیم

۷- " " " حداقل " " " " "

x	0	1	2
$P(x)$	%20	%50	%30

y	0	1	2	3
$P(y)$	%10	%30	%25	%25

۱-

۲- %10

$F(1,1) = \%10$ (صبح و عصر یک غایب)

۳- %20

$P(x>y) = P(1,0) + P(2,0) + P(2,1) = \%20$

یعنی ۸۰٪ درصد از روزها، غایبین عصر از صبح بیشتر باشد

۴- %25

$P(x=y) = P(0,0) + P(1,1) + P(2,2) = 0/25$

۵-

$Z = x+y$	0	1	2	3	4	5
	%5	%10	%20	%40	%20	%5
		$(0,1)$	$(2,0)$	$(2,1)$	$(1,3)$	$(2,3)$
		$(1,0)$	$(0,2)$	$(1,2)$	$(2,2)$	
			$(1,1)$	$(0,3)$		

$= \%100$

۶-

$P(x+y \leq 3)$ حداکثر

$P(x+y=0) + P(x+y=1) + P(x+y=2) + P(x+y=3)$

از جدول بالا $\Rightarrow$ 0, 1, 2, 3

$\%5 + \%10 + \%20 + \%40 = \%75$

۷-

حداقل

$P(x+y \geq 3)$ $\%40 + \%20 + \%5 = \%65$

Subject .... Date ....

کوواریانس : شاید تعریف هم سوال بیاد مهم

معیاری است عددی برای بررسی میزان تاثیر دو متغیر بر روی یکدیگر. به عبارت دیگر می‌خواهیم بدانیم افزایش یا کاهش یک متغیر بر روی متغیر دیگر تاثیر مستقیم یا معکوس دارد یا خیر.

$$Cov(x,y) = E(xy) - E(x)E(y)$$

توزیع حاشیه‌ای x $\quad E(x) = \sum x\,p(x)$

y " " $\quad E(y) = \sum y\,p(y)$

" توام $\quad E(x,y) = \sum xy\,p(xy)$

مثال : در مثال قبل (آزمایش صبح و عصر) مطلوب است محاسبه کوواریانس؟

$0 \times 0 \times ۰/۰۵ = 0 \leftarrow xy$ برابر

x \ y	۰	۱	۲	۳	
۰	۰/۰۵	۰/۰۵	۰/۱۰	۰	۰/۲۰
۱	۰/۰۵	۰/۱۰	۰/۲۵	۰/۱۰	۰/۵۰
۲	۰	۰/۱۵	۰/۱۰	۰/۰۵	۰/۳۰

$2 \times 2 \times ۰/۱۰ \rightarrow$ برابر xy اینطوری محاسبه میشود

مقدار $= xy\,p(xy) = 0$ (۰ ، ۰ ، ۰/۰۵)

x	۰	۱	۲
P(x)	۰/۲۰	۰/۵۰	۰/۳۰
$E(x)=\sum x p(x)$	۰	۰/۵۰	۰/۶۰

$\Rightarrow E(x) = 1/1$

y	۰	۱	۲	۳
P(y)	۰/۱۰	۰/۳۰	۰/۴۵	۰/۱۵
$E(y) = \sum y p(y)$	۰×۰/۱۰	۱×۰/۳۰	۲×۰/۴۵	۳×۰/۱۵

$\Rightarrow E(y) = 1/65$

$$\sum xy\,p(xy) = ۰/۱۰ + ۰/۵۰ + ۰/۳۰ + ۰/۳۰ + ۰/۴۰ + ۰/۳۰ = ۱/۹$$

که این اعداد از ضرب اعداد سطر و ستون و عدد درصد بدست آمده که باهم جمع کنیم

$$cov(x,y) = E(xy) - E(x)E(y)$$
$$= ۱/۹ - (۱/۱) \times (۱/۶۵) = ۰/۰۸۵$$

Corr

ضریب همبستگی: (کورولیشن) (تعریف کلمات) کورولیشن

معیاری است برای بررسی میزان همبستگی دو متغیر که برخلاف کوواریانس به مقیاس عددی ربط ندارد. بین ۱ و ۱- می‌باشد.

$$-1 \leq Corr \leq 1$$

اگر ضریب همبستگی دو متغیر نزدیک به عدد ۱ باشد یعنی دو متغیر به یکدیگر وابسته هستند و کاهش اولی منجر به کاهش دومی شده و بالعکس. زمانی که ضریب همبستگی به عدد ۱- نزدیک می‌شود دو متغیر به یکدیگر وابسته اند اما در خلاف جهت یعنی افزایش یکی باعث کاهش دیگری می‌گردد و بالعکس و در حالتی که ضریب همبستگی نزدیک صفر می‌باشد دو متغیر ارتباطی به یکدیگر ندارند.

$$Corr = \frac{Cov(x,y)}{\sigma_x \cdot \sigma_y}$$

انحراف معیار

مهم

$$\sigma_x^2 = E(x^2) - E^2(x) \Rightarrow \sigma_x = \sqrt{\sigma_x^2}$$

$$\sigma_y^2 = E(y^2) - E^2(y) \Rightarrow \sigma_y = \sqrt{\sigma_y^2}$$

$$E(x^2) = \sum x^2 p(x)$$

$$E(y^2) = \sum y^2 p(y)$$

$$E^2(x) = (E(x))^2$$

$$E^2(y) = (E(y))^2$$

مثال: در مثال قبل مطلوب است محاسبه ضریب همبستگی (corr) و تفسیر آن؟

$Cov(x,y) = 0.085$

$E(x) = 1.1 \Rightarrow E^2(x) = 1.21$

$E(y) = 1.65 \Rightarrow E^2(y) = 2.72$

از مثال قبل

برای $E(x^2)$ و $E(y^2)$ در جدول صفحه قبل محاسبه می‌کنیم.

$$E(x^2) = x^2 p(x) = 0^2 \times 0.20 + 1^2 \times 0.5 + 2^2 \times 0.30 = \boxed{1.7}$$ $E(x^2)$

$$E(y^2) = y^2 p(y) = 0^2 \times 0.10 + 1^2 \times 0.3 + 2^2 \times 0.45 + 3^2 \times 0.15 = \boxed{3.45}$$ $E(y^2)$

$$\sigma_x^2 = 1.7 - (1.1)^2 = 0.49 \Rightarrow \sigma_x = \sqrt{0.49} = \boxed{0.7}$$

$$\sigma_y^2 = 3.45 - (1.65)^2 = 0.7275 \Rightarrow \sigma_y = \sqrt{0.7275} = \boxed{0.85}$$

$$Corr = \frac{0.085}{0.7 \times 0.85} = \boxed{0.14}$$ ⟶ ضعیف / نزدیک به صفر می باشد پس نتیجه می گیریم ناچیز یعنی تقریباً هیچ ارتباطی به هم ندارند (دو متغیر به یکدیگر ارتباطی ندارند)

استقلال دو متغیر تصادفی:

نقطه‌ی مقابل همبستگی و کوواریانس استقلال دو متغیر به معنی رخ دادن یا رخ دادن آن به متغیر دیگر ارتباطی ندارد

$E(xy) = E(x)\,E(y)$ شرط استقلال

نکته: اگر x و y مستقل باشند $\Longrightarrow$ $Corr(x,y) = 0$ و $Cov(x,y) = 0$

عکس این نکته صحیح نمی باشد.

تمرین: نکته فوق را اثبات کنید.

$$E(xy) = E(x)E(y) \Rightarrow Cov = \cancel{E(xy) - E(xy)} = 0$$

یا $Cov(x,y) = E(xy) - E(x)E(y)$

$$E(xy) = E(x)E(y)$$
$$E(xy) - E(x)E(y) = 0$$

Subject .... Date ....

مثال: سکه‌ای قرار است سه بار پرتاب شود. فرض کنید x تعداد شیرها در پرتاب‌های اول و دوم و y برابر با تعداد شیرها در پرتاب سوم.

الف) توزیع احتمال توام آن را بنویسید.

بررسی کنید که آیا x و y مستقل هستند یا خیر.

$x \backslash y$	0	1	
0	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{2}{8}$
1	$\frac{2}{8}$	$\frac{2}{8}$	$\frac{4}{8}$
2	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{2}{8}$
	$\frac{4}{8}$	$\frac{4}{8}$	$\frac{8}{8}$

$x = 2, 1$ تعداد شیرها در پرتاب

$y = 3$ " " " "

$2^3 = 8$ حالت دارد

شیر T خط H

یعنی در دو پرتاب اول شیر باشد در پرتاب سوم هم شیر باشد: T T T $\rightarrow F(2,1) = \frac{1}{8}$

T T H $\rightarrow F(2,0) = \frac{1}{8}$

یعنی در پرتاب اول و دوم یکی شیر باشد و در پرتاب سوم شیر نباشد: T H T $\rightarrow F(1,0) = \frac{2}{8}$

x	0	1	2	
$p(x)$	$\frac{2}{8}$	$\frac{4}{8}$	$\frac{2}{8}$	
$\sum x\,p(x)$	0	$\frac{4}{8}$	$\frac{4}{8}$	$\frac{8}{8}=1$

$E(x) = 1$

y	0	1	
$p(y)$	$\frac{4}{8}$	$\frac{4}{8}$	
$\sum y\,p(y)$	0	$\frac{4}{8}$	$\frac{4}{8}$

$E(y) = \frac{4}{8}$

شرط استقلال $\Leftarrow E(xy) = E(x)\,E(y)$

$E(xy) = \frac{4}{8}$

$E(x) = 1$

$E(y) = \frac{4}{8}$

$\Rightarrow \frac{4}{8} = 1 \times \frac{4}{8} = \frac{4}{8}$

۲ متغیر مستقلند

Scanned by CamScanner

Subject .... Date ....

آزمون فرض (تئوری رفع مجهم است)

حدسی درباره‌ی خصوصیتی از جامعه که قویاً قابل تایید است یا خیر معمولاً این حدس در مورد پارامترهای جامعه مطرح می‌شود.
در واقع هر حکمی درباره‌ی جامعه یک آزمون فرض آماری است.

H_0	ادعا صحیح است	H_0	فرض صفر
H_1	〃 〃 نیست	H_1	〃 مقابل

در علم آمار بنا بر پذیرش فرض صفر است و فرض صفر زمانی رد می‌شود که فرض مقابل (H_1) قویاً قابل اثبات باشد.

در علم ریاضی قضایا (مسائل) ریاضی با استناد به ادله موجود یا رد می‌شوند یا اثبات، مثلاً جواب‌های یک معادله درجه دوم کاملاً مشخص است (درست یا غلط) اما در نقطه مقابل در علم آمار عدم قطعیت نسبت به آزمون فرض آماری وجود دارد (شک نسبت به نتایج)

خطای نوع اول و خطای نوع دوم:

خطای نوع اول α
〃 نوع دوم β

رد کردن H_0 وقتی که H_0 درست است: خطای نوع اول
پذیرفتن H_0 وقتی که H_1 درست است: 〃 〃 دوم

جدول خطاها:

	H_0 صحیح (درست)	H_0 غلط
H_0 رد نمی‌شود	صحیح	β خطای نوع دوم
H_0 رد می‌شود	خطای نوع اول α	صحیح

فرض غلط را اشتباه بپذیریم

مهم است.

فرض درست را به غلط رد کنیم

رگرسیون:

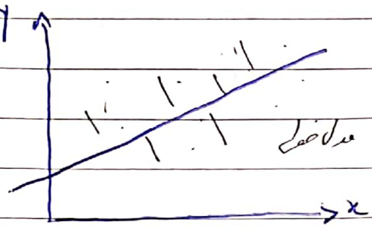
یک فن آماری است که برای مدل‌سازی و ارتباط بین متغیرها استفاده می‌شود. ابتدا تحلیل‌گر حدسی در مورد ارتباط بین دو متغیر می‌زند با این استنباط که این ارتباط، یک ارتباط خطی است در صورتی که

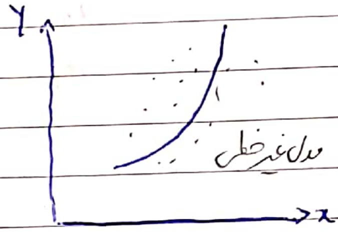
Subject .... Date ....

هرگاه داده‌ها تقریباً در امتداد یک خط باشند، این حدس تحلیل‌گر تایید می‌شود. در این صورت معادله‌ی خط رگرسیون به صورت زیر خواهد بود. فرمول خط رگرسیون

$$y = \alpha + \beta x$$

عرض از مبدا — دبا — شیب (ضریب رگرسیون)





$$\beta = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2} \qquad \beta = \frac{\sum xy - n\bar{x}\bar{y}}{\sum x^2 - n\bar{x}^2}$$

$$\beta = \frac{cov(xy)}{\sigma_x^2} \qquad r = \frac{\sum xy - n\bar{x}\bar{y}}{\sqrt{\sum x^2 - n\bar{x}^2}\sqrt{\sum y^2 - n\bar{y}^2}}$$

واریانس

کوواریانس

$$\boxed{\alpha = \bar{y} - \beta\bar{x}}$$

مثال، اطلاعات زیر در دست است. شیب خط رگرسیون را بدست آورید. عرض از مبدا آن را بنویسید. معادله‌ی خط را بنویسید و در صورت امکان خط آن را رسم کنید.

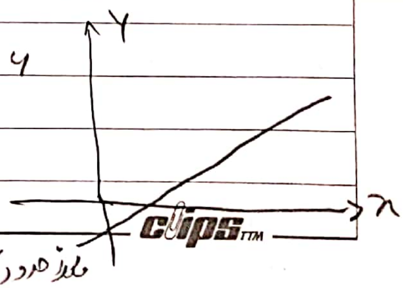
$$\sum x = 13 \quad \sum y = 27 \quad \sum xy = 108 \quad \sum x^2 = 51 \quad n = 5$$

$$\beta = \frac{n\sum xy - \sum x\sum y}{n\sum x^2 - (\sum x)^2} = \frac{5\times 108 - 13\times 27}{5\times 51 - 13^2} = \frac{189}{86} = \boxed{2.19 = \beta}$$

$$\alpha = \bar{y} - \beta\bar{x} = 5.4 - (2.19)(2.6) = -0.2 \qquad \boxed{\alpha = -0.2}$$

$$\bar{y} = \frac{\sum y}{n} = \frac{27}{5} = 5.4 \qquad \bar{x} = \frac{\sum x}{n} = \frac{13}{5} = 2.6$$

$$y = -0.2 + 2.19x \Rightarrow y = \boxed{2.19x - 0.2}$$



مقدار عرض از مبدا ۰٫۲-

Scanned by CamScanner

Subject .... Date ....

مثال: اگر $x=6$ باشد y را بیابید.

$$\bar{y} = \frac{\sum y}{n} \Rightarrow 10 = \frac{\sum y}{20} \Rightarrow \sum y = 200$$

$\sum xy = 1150$

$\sum x^2 = 550$

$\bar{x} = 5$

$\bar{y} = 10$

$n = 20$

$$\beta = \frac{\sum xy - n\bar{x}\bar{y}}{\sum x^2 - n\bar{x}^2} = \frac{1150 - 20 \times 5 \times 10}{550 - 20 \times 5^2} = \boxed{3 = \beta}$$

$$\alpha = \bar{y} - \beta\bar{x} = 10 - 3 \times 5 = -5 \Rightarrow \alpha = -5$$

$$\left.\begin{array}{l} y = 3x - 5 \\ x = 6 \end{array}\right\} \Rightarrow y = 18 - 5 = 13 \qquad y = 13$$

صورت مسئله