SUBJECT:
Year ( ) Month ( ) Date ( )

رزمخواه

سرفصل آمار و احتمال — منابع — (بیودیان x)

۱- احتمال — ۱- مقدمه‌ای بر آمار و احتمال برای مهندسان و محققان علوم

۲- متغیرهای تصادفی — تألیف: والپل — ترجمه: دکتر مجید اسدی و ابوالقاسم بزرگ‌نیا

۳- توزیع‌های خاص احتمالی — ۲- مقدمه‌ای بر آمار ریاضی — تألیف والپول و فروند ← پیچیده‌تر

۴- نمونه‌گیری و برآورد — ۳- مفاهیم و روش‌های آماری (هر دو جلد)

۵- آزمون فرضیه‌های آماری — تألیف: باتاچاریا — ترجمه: مرتضی ابن‌شهرآشوب و فتاح میکائیلی

میان ترم ۸-۶ ن هفته آخر فروردین (یکشنبه) — ۴- آمار و کاربرد آن در مدیریت ← مسئله‌های زیاد و ساده

پایان ترم ۱۴-۱۲ ن — تألیف: دکتر عادل آذر

← تمرین‌های کلاس مهم‌اند!

فصل اول:

آزمایش تصادفی: آزمایشی است که نتیجه‌ی آن از قبل معلوم نباشد.

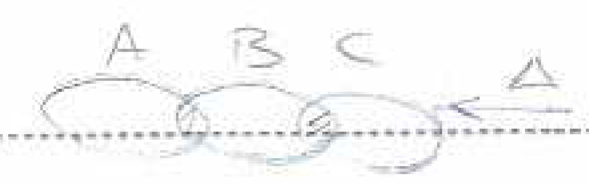
فضای نمونه‌ای آزمایش: به مجموعه‌ی کلیه‌ی حالات ممکن رخداد یک آزمایش تصادفی، فضای نمونه‌ی آن آزمایش گفته می‌شود و آن را معمولاً با S نشان می‌دهیم. مثلاً در آزمایش پرتاب سکه: $S=\{\text{خط}, \text{شیر}\}$

پیشامد تصادفی: به هر یک از زیرمجموعه‌های فضای نمونه‌ی یک آزمایش تصادفی گفته می‌شود. 2^2 (2^n)

SUBJECT:
Year ( ) Month ( ) Date (

اعمال روی پیشامدها: متمم، اجتماع، اشتراک، تفاضل، تقارن

پیشامدهای ناسازگار: دو پیشامد A و B را که اشتراکی نداشته باشند ناسازگار می‌گوییم یعنی:

$A\cap B=\emptyset$ ← سه پیشامد ناسازگار ... 

* سه « دو به دو « ...

افراز فضای نمونه: مجموعه‌های A_1، A_2، ...، A_n را یک افراز برای فضای نمونه S می‌گوییم هرگاه:

الف) $S=\bigcup_{i=1}^{n} A_i$ ب) A_iها دو به دو ناسازگار باشند یعنی $A_i\cap A_j=\emptyset \; ; i\neq j$

قوانین احتمال:

$P(A)=\dfrac{n(A)}{n(S)}$ ← به شرط همشانسی ... یکنواخت

طول
سطح
حجم
فضای نمونه‌ای

1) $\forall A\in S \; ; \; 0\le P(A)\le 1$ (تعریف احتمال)

2) $P(\emptyset)=0$

نکته. توجه شود که هر پیشامدی با احتمال صفر لزوماً تهی نیست. $P(A)=0 \not\Rightarrow A=\emptyset$

مثال. احتمال رخداد یک نقطه وقتی فضای نمونه بازه‌ای از اعداد حقیقی باشد (یعنی طول داشته باشد) صفر است.

3) پیشامد قطعی (حتمی) $P(S)=1$

4) $P(A\cup B)=P(A)+P(B)-P(A\cap B)$

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

SUBJECT: ۲

Year (   ) Month (   ) Date (   )

۶) اگر $A_1, \dots, A_n$ پیشامدهای دوبه‌دو ناسازگار باشند آنگاه: $P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i)$

۷) $P(A^c) = 1 - P(A)$

۸) $P(A-B) = P(A) - P(A\cap B)$ $A-B = A\cap B'$ اثبات

اثبات: $A = \underbrace{(A-B)}\cup\underbrace{(A\cap B)} \longrightarrow \dots$

ناسازگارند

جلسه دوم:

تعریف احتمال: احتمال عددی است بین ۰ و ۱ که میزان عدم قطعیت رخداد پیشامدها را مشخص می‌کند.

اگر فضای نمونه‌ی آزمایش تصادفی مجموعه‌ای از اعداد یا اشیا باشد، آنگاه احتمال رخداد پیشامد آن به‌صورت

زیر محاسبه می‌شود: $\star\ P(A) = \frac{n(A)}{n(S)}$

که در آن منظور از $n(A)$ $\star$ تعداد اعضای پیشامد A است.

اگر فضای نمونه بازه‌ای از اعداد حقیقی باشد، آنگاه احتمال رخداد A می‌شود: $P(A) = \frac{A \text{ طول}}{S \text{ طول}}$

قانون احتمال شرطی: اگر از رخداد پیشامدی مانند B مطلع باشیم و بخواهیم احتمال رخداد پیشامد دیگری

را محاسبه کنیم از فرمول شرطی زیر استفاده می‌کنیم. $P(A|B) \overset{\star}{=} \frac{P(A\cap B)}{P(B)}$

$P(A|B) = \frac{n(A\cap B)/n(S)}{n(B)/n(S)} \overset{\star}{=} \frac{n(A\cap B)}{n(B)}$ مثل اینکه فضای نمونه B باشد!

SUBJECT:

Year ( ) Month ( ) Date (

مثال. اگر بدانیم در آزمایش پرتاب یک تاس خال رو شده زوج است، چه قدر احتمال دارد آن خال ۲ باشد؟ چه قدر احتمال دارد آن خال ۳ باشد؟ ۱/۳ ، صفر

قانون ضرب احتمال. برای هر دو پیشامد A و B داریم:

$$P(A\cap B) = P(B)\,P(A|B)$$

$$P(A\cap B) = P(A)\,P(B|A)$$

همچنین برای سه پیشامد A، B و C می‌توان نوشت:

$$P(A\cap B\cap C) = P(A)\,P(B|A)\,P(C|A\cap B)$$

$$= P(B)\,P(C|B)\,P(A|B\cap C)$$

مثال. از ظرفی محتوی ۴ مهره‌ی سفید، ۶ مهره‌ی سیاه و ۵ مهره‌ی قرمز است، ۳ مهره به تصادف بدون جایگذاری خارج می‌کنیم چه قدر احتمال دارد هر سه مهره سفید باشد؟

ب- مهره‌ها به ترتیب سفید، سیاه و قرمز باشند؟

حل الف. پیشامد اینکه مهره‌ی i ام سفید باشد A_i ، $i=1,2,3$

$$P(A_1\cap A_2\cap A_3) = P(A_1)\,P(A_2|A_1)\,P(A_3|A_1\cap A_2) = \frac{4}{15}\times\frac{3}{14}\times\frac{2}{13}$$

حل ب. پیشامد اینکه مهره‌ی i ام سفید باشد A_i ، $i=1,2,3$

B_i « سیاه « « « «

C_i « قرمز « « « «

$P(A_1 \cap B_2 \cap C_3) = P(A_1) P(B_2 | A_1) P(C_3 | A_1 \cap B_2) \overset{m.}{=} \frac{4}{15} \times \frac{6}{14} \times \frac{5}{13}$

استقلال پیشامدهای تصادفی.

دو پیشامد A و B را مستقل گوییم هرگاه رخداد یکی بر دیگری تاثیری نداشته باشد یعنی $P(A|B) = P(A)$

$P(B|A) = P(B)$

به عبارتی A و B را مستقل گوییم اگر و فقط اگر: $P(A \cap B) = P(A) P(B)$

نکته: پیشامدهای A_1، ...، A_n را مستقل گوییم اگر و تنها اگر: $P(\bigcap_{i=1}^{n} A_i) = \prod_{i=1}^{n} P(A_i)$

★ A و B و C دو به دو مستقل:

$P(A \cap B) = P(A) P(B)$ $P(A \cap C) = P(A) P(C)$

$P(B \cap C) = P(B) P(C) \xleftarrow{?} ★ \; P(A \cap B \cap C) = P(A) P(B) P(C)$

قانون احتمال کل.

فرض کنید A_1، ...، A_n یک افراز از فضای نمونه S باشد و B نیز پیشامدی دلخواه در S باشد، در اینصورت:

$B = (B \cap A_1) \cup (B \cap A_2) \cup \cdots (B \cap A_n) \Rightarrow P(B) = \sum_{i=1}^{n} \underbrace{P(B \cap A_i)}_{P(B|A_i) P(A_i)}$

$P(B) = \sum_{i=1}^{n} P(B|A_i) . P(A_i)$ قانون احتمال کل.

قانون بیز: برای دو پیشامد A و B قانون بیز بیان می‌کند:

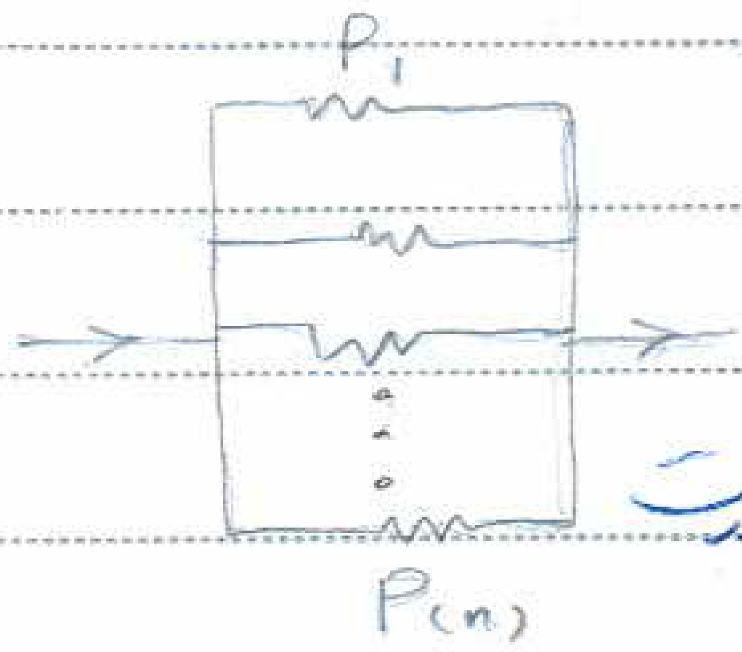
SUBJECT:
Year ( ) Month ( ) Date (

$$P(A|B) = \frac{P(A\cap B)}{P(B)} \overset{*}{=} \frac{P(A)P(B|A)}{P(B)}$$

اگر A_1 ، ... ، A_n افرازی از فضای نمونه باشند، قانون بیز را می‌توان به صورت زیر نوشت:

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{j=1}^{n} P(B|A_j)P(A_j)} \qquad i = 1,2,\dots,n$$

مثال. یک سیستم موازی در نظر بگیرید که احتمال فعال بودن هر یک از مولفه‌های آن برابر P_i (که $i=1,\dots,n$) باشد، چه قدر احتمال دارد سیستم عمل کند؟



حل. می‌دانیم در یک سیستم موازی مولفه‌ها مستقل از هم عمل می‌کنند، احتمال مطلوب به صورت زیر محاسبه می‌شود:

$$P(\underset{x\geq 1}{\text{حداقل یکی از مولفه‌ها سالم باشند}}) \overset{*}{=} 1 - P(\text{هیچ کدام سالم نباشند}) \qquad x<1 \equiv x=0$$

پیشامد اینکه مولفه i ام سالم باشد: A_i

$$= 1 - P\left(\bigcap_{i=1}^{n} A_i^c\right) =$$

* A_i ها مستقل اند پس A_i^c ها هم مستقل اند

$$= 1 - \prod_{i=1}^{n} P(A_i^c) = 1 - \prod_{i=1}^{n} (1 - \underbrace{P(A_i)}_{P_i}) = 1 - \prod_{i=1}^{n} (1-P_i) \neq 1-(1-P_i)^n$$

حل دوم. (بحث نشد...) $P\left(\bigcup_{i=1}^{n} A_i\right)$

نکته. اگر A ، B مستقل باشند آنگاه:

B, A' B', A B', A' نیز مستقل هستند.

SUBJECT: ۵

Year ( ) Month ( ) Date ( )

حل:

$S = \{$ (شیر،شیر) ، خط خط ، شیر خط ، خط شیر $\}$

$X:$ ۲ ، ۰ ، ۱ ، ۱

$\Rightarrow R_X = \{0, 1, 2\}$

$$P(X=x) = \begin{cases} 1/4 & x=0 \\ 1/2 & x=1 \\ 1/4 & x=2 \end{cases}$$

★ تعریف. به تابع احتمال $P(X=x)$ تابع جرم احتمال متغیر تصادفی گسسته X می‌گوییم اگر:

الف) $0 \le P(X=x) \le 1 \quad \forall x \in R_X$

ب) $\sum_{x \in R_X} P(X=x) = 1$

امید ریاضی : (متوسط ...)

مقدار مورد انتظار $\bar{X}$ / امید ریاضی

$\boxed{\text{X بیانگر یک جامعه است}}$

فرض کنید X یک متغیر تصادفی گسسته با تابع جرم احتمال $P(X=x)$ باشد، در اینصورت ، امید ریاضی X عبارتست از : تعداد +

$$E(X)=\sum_{x\in R_X} x\,\underbrace{P(X=x)}_{w_i}$$

کل جامعه

ملاحظه می شود که امید ریاضی متغیر تصادفی X که در بالا تعریف کردیم میانگین وزنی مقادیر متغیر تصادفی X است .

$$\begin{matrix} x_1 & \cdots & x_n \\ \downarrow & & \downarrow \\ w_1 & & w_n \end{matrix} \quad , \sum w_i = 1 \longrightarrow \bar{x}_w = \sum_{1}^{n} w_i x_i$$

* نکته: امید ریاضی X یک مقدار ثابت است و به x بستگی ندارد .

مثال: در آزمایش دوبار پرتاب یک سکه متوسط (امید ریاضی) تعداد شیرها را حساب کنید .

یعنی باید ابتدا X تعریف کنیم ...

S={خ خ ، ش خ ، خ ش ، ش ش}

حل: X تعداد شیرها در دوبار پرتاب سکه پس : $R_X=\{0,1,2\}$

$$P(X=x)=\begin{cases} 1/4 & x=2 \text{ یا } 0 \\ 1/2 & x=1 \\ (\text{a}\circ & \text{aw})? \end{cases}$$

x	0	1	2
P(X=x)	1/4	1/2	1/4

$$\Rightarrow E(X)=0\times\frac{1}{4}+1\times\frac{1}{2}+2\times\frac{1}{4}=1$$

(برای تعداد دفعات بیشتر E(x) که تکرار شد)

تمرین - در مثال قبل فرض کنید Y درآمد فرد A در یک بازی باشد که به طریق زیر مطرح می شود :
در دوبار پرتاب یک سکه اگر هر دوبار شیر ظاهر شود شخص A ، ۱۰ تومان از شخص B بگیرد ، اگر یک خط یا دو خط ظاهر شود به ترتیب ۲ تومان و ۵ تومان به شخص B بدهد .

SUBJECT: ۴
Year ( ) Month ( ) Date ( )

نکته ۲ : اگر A_1، ... ، A_n پیشامدهای دو به دو مستقل باشند آنگاه هر ترکیبی از چند تای آن ها از هر ترکیبی دیگر از مابقی آن ها نیز مستقل است.

نتیجه : اگر A_1، ... ، A_n مستقل باشند آنگاه A_1'، ... ، A_n' هم از هم مستقل هستند.

تمرین : فرض کنیم ترکیبی دارای سه کارخانه است که میزان درصد تولید و کالاهای معیوب هر کارخانه در جدول زیر آمده است.

	درصد تولید	درصد معیوب
کارخانه ی ۱	۳۰٪	۲٪
۲ "	۴۵٪	۳٪ → P(تعلق به کارخانه ی ۲ | معیوب)
۳ "	۲۵٪	۴٪

☆ اگر کالاهای معیوب خریداری شود چه قدر احتمال دارد از دومی باشد؟ (با قانون بیز) book

حل : اگر A پیشامد معیوب بودن کالا باشد و B_i پیشامد تعلق کالا به کارخانه ی i باشد:

$$P(B_2|A) = \frac{P(A|B_2)P(B_2)}{\sum_{j=1}^{3} P(A|B_j)P(B_j)} \overset{☆}{=} \frac{0/03 \times 0/45}{0/02 \times 0/30 + 0/03 \times 0/45 + 0/04 \times 0/25} =$$

(نه ...)

$\frac{27}{135}$ $60 + 225 = 295$ 59

$$= 0/3417 = 34/2\ \%$$

فصل دوم : متغیرهای تصادفی

متغیر تصادفی : عملگری است که فضای نمونه ی آزمایش تصادفی را طبق تعریفی خاص به زیرمجموعه ای از اعداد حقیقی می برد، به عبارتی یک متغیر تصادفی با نماد X تابعی است از S (دامنه ی تابع) به زیرمجموعه ای از اعداد حقیقی مانند R_X (برد تابع).

$$X : S \longmapsto R_X \subseteq R$$

به R_X تکیه گاه متغیر تصادفی X گفته می شود.

اگر R_X مجموعه ای شمارا باشد به X یک متغیر تصادفی گسسته و در غیر اینصورت به X یک متغیر تصادفی پیوسته می گوییم.

مثال : اگر $R_X = \{0, 1, 2\}$ باشد، X متغیر تصادفی گسسته و ...

$S = \{$خط , شیر$\}$

$S = \{$خخ , شخ , خش , شش$\}$ → 0 , 1 , 1 , 2

اگر $R_X = (0, 1)$ باشد، X متغیر تصادفی پیوسته است.

تابع جرم احتمال : یک متغیر تصادفی گسسته، هر یک از مقادیر تکیه گاه خود را با احتمال مشخص اختیار می کند.

به مجموعه ی این احتمالات تابع جرم احتمال آن متغیر تصادفی گفته می شود. $P(X = x)$

مثال : در آزمایش دو بار پرتاب یک سکه فرض کنید X نشانگر تعداد شیرها باشد. تابع جرم احتمال X را حساب کنید.

SUBJECT: ✓
Year ( ) Month ( ) Date ( )

متوسط درآمد شخص A را در هر بازی حساب کنید. * (در چنین بازی اگر $E(x)$ صفر شود بازی عادلانه است)

تمرین ۲. فرض کنید درآمد شخص A در یک بازی دارای تابع جرم احتمال زیر باشد.

x	c	-2	10
$P(X=x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

مقدار c را طوری تعیین کنید که این بازی یک بازی عادلانه باشد؟

نکته: منظور از بازی عادلانه بازی است که متوسط درآمد در آن صفر باشد.

حل تمرین ۱.

$S = \{$ش ش، خ ش، ش خ، خ خ$\}$

$X:\ 10 \quad -2 \quad -2 \quad -5$

X: درآمد

x	10	-2	-5
$P(X=x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

$$\Rightarrow E(x) = \frac{10}{4} - \frac{2}{2} - \frac{5}{4} = 0.25$$

$1000 \times 0.25 = 250$ (درآمد) ... تکرار از

تمرین ۲)

$$E(x) = 0 \Rightarrow E(x) = -2 \times \frac{1}{2} + 10 \times \frac{1}{4} + c \times \frac{1}{4} = 0$$

$$\Rightarrow c = -6$$

SUBJECT:

Year ( ) Month ( ) Date (

واریانس: فرض کنید X دارای تابع جرم احتمال $P(X=x)$ باشد، دراینصورت واریانس x عبارت است از: (پراکنش: بیان اندازه پراکندگی)

$$\sigma_x^2 = Var(X) = E(X - E_X)^2$$

$\underbrace{E_X \text{ چون ثابت باید}}$ $\Rightarrow Var(X) = E(X^2) - (EX)^2$

نکته ۱: واریانس یک متغیر تصادفی یک مقدار ثابت است.

$E(X^2 + E_X^2 - 2XE_X) = EX^2 - 2EXEX + (EX)^2 = EX^2 - (EX)^2$

خواص امید ریاضی و واریانس:

مقدار ثابت

1) $E(a) = a$ 1) $Var(a) = 0$

2) $E(ax) = aE(x)$ 2) $Var(X) \geq 0$

3) $E(ax+b) = aE(x) + b$ 3) $Var(aX) = a^2 Var(x)$

4) $E(ax+bY+C) = aE(x) + bE(y) + c$ 4) $Var(ax+b) = a^2 Var(X)$

تعریف انحراف معیار. انحراف معیار متغیر تصادفی X به صورت زیر تعریف می‌شود:

$$\sigma_x = \sqrt{\sigma_x^2} = \sqrt{Var(x)}$$

تمرین. در مثال ٤ ی قبل واریانس متغیر مورد مطالعه را حساب کنید.

نکته. فرض کنید X یک متغیر تصادفی با تابع جرم احتمال $P(X=x)$ و g یک تابع حقیقی باشد دراینصورت:

برای Var و ...

$$E(g(x)) = \sum_{x \in R_x} g(x) P(X=x)$$

STAEDTLER

book: محاسبات دیگری نسبت به Var ؟

- مثلاً برای $g(x)=x^2$ داریم:
$$E(x^2)=\sum_{x\in R_x} x^2 P(X=x)$$

در مثال ۱:
$$E(x^2)=0\times 1/4+1\times 1/2+2^2\times 1/4=1/5$$

ت. Var باید مثبت باشد.
$$\Rightarrow Var(x)=E(x^2)-(Ex)^2=1/5-1=0/5$$

تابع توزیع احتمال (تابع احتمال تجمعی / انباشته)

فرض کنید X یک متغیر تصادفی باشد در این صورت تابع توزیع احتمال X در نقطه‌ی t به صورت زیر محاسبه می‌شود:
$$F_X(t)=P(X\le t)=\sum_{x\le t} P(X=x)$$

مثال: در مثال ۱، تابع احتمال تجمعی X را حساب کنید. (دو بار پرتاب سکه ...)

$$F_X(t)=\begin{cases}0 & t<0\\ 1/4 & 0\le t<1\\ 3/4 & 1\le t<2\\ 1 & t\ge 2\end{cases}$$

مثال ت. $\Rightarrow$ $\begin{cases}F_X(10)=1\\ F_{2/5}=1\\ F_{1/8}=3/4\end{cases}$

خواص تابع توزیع احتمال:

۱) تابع توزیع احتمال یک تابع غیرنزولی است.

۲) تابع توزیع احتمال در هر نقطه حداقل از راست پیوسته است.

۳) $F(+\infty)=\lim_{x\to+\infty} F(x)=1$

۴) $F(-\infty) = \lim_{x\to -\infty} F(x) = 0$

نکته: اگر X گسسته باشد آنگاه $P(X=x) = F_X(x) - F_X(x^-)$ (معنی *) که در آن x^- نقطه‌ی قبل از x در تکیه‌گاه متغیر تصادفی X است.

* تمرین: محاسبه‌ی واریانس: ۱) در مثال تعداد شیر در دو بار پرتاب سکه:

راه ۱) $Var(x) = E(x-Ex)^2 = E(x-1)^2 = \sum_{x\in R_x} g(x)P(X=x) =$ (* $g(x)$)

$= 1\times \frac{1}{4} + 0 + 1\times\frac{1}{4} = \frac{1}{2}$

راه ۲) $Var(x) = E(x^2) - (Ex)^2 = 1\times\frac{1}{2} + (2)^2\times\frac{1}{4} + 0 - 1 = \frac{1}{2}$ ✓

۲) مثال بازی: راه ۱) $Var(x) = E(x-\frac{1}{4})^2 = (\frac{39}{4})^2\times\frac{1}{4} + (\frac{-9}{4})^2\times\frac{1}{2} + (\frac{-21}{4})^2\times\frac{1}{4}$

$= 33.1875$

راه ۲) $Var(x) = E(x^2) - (Ex)^2 = \frac{100}{4} + \frac{4}{2} + \frac{25}{4} - \frac{1}{16} = 33.1875$ ✓

۳) مثال بازی عادلانه: راه ۱، ۲) $Var(x) = E(x-0)^2 = E(x^2) - (\overset{0}{Ex})^2 = \frac{36}{4} + \frac{4}{2} + \frac{100}{4} = 36$

(book) تذکر: همچنین (راه ۲): $Var(x) = E(x^2) - (E_x)^2$ چون:

$Var(x) = E(x-E_x)^2 = E(x^2 + \underset{\text{عدد ثابت}}{E_x^2} - 2xE_x) \xrightarrow[\text{امید ریاضی}]{\text{طبق خواص}} E(x^2) + (E_x)^2 - 2E_xE_x$

$= E(x^2) - (E_x)^2$ ✓

SUBJECT: ۹

Year ( ) Month ( ) Date ( )

تابع چگالی احتمال:

تعریف: به تابع $f(x)$ تابع چگالی احتمال متغیر تصادفی پیوسته X گوییم هرگاه:

۱) $\forall x ; \; f(x) \geq 0$

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

نکته: اگر X یک متغیر تصادفی پیوسته با تابع چگالی احتمال $f(x)$ باشد و a, b مقادیر ثابتی باشند آنگاه:

الف) $P(X=\alpha) = 0$

ب) $P(a < x < b) = \int_a^b f(x)\,dx$

امید ریاضی و واریانس.

فرض کنید X یک متغیر تصادفی پیوسته با تابع چگالی احتمال $f(x)$ باشد در اینصورت:

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

یک تابع حقیقی

$Ex^2 = \int_{-\infty}^{+\infty} x^2 f(x)\,dx$

$Var(x) = E(x - E_x)^2 = E(x^2) - (E_x)^2$

STAEDTLER

SUBJECT:

Year ( ) Month ( ) Date (

تابع توزیع احتمال:

اگر x یک متغیر تصادفی پیوسته با تابع چگالی احتمال $f(x)$ باشد آنگاه:

$$F_X(y) = P(X \le y) = \int_{-\infty}^{y} f(x)\,dx$$

نکته: اگر X (متغیر پیوسته‌ی تصادفی) دارای تابع توزیع احتمال $F_X(x)$ باشد، آنگاه:

$$f_X(x) = \frac{d}{dx} F_X(x)$$

مثال:

فرض کنیم طول عمر یک قطعه‌ی الکترونیکی (X) دارای تابع توزیع احتمال زیر باشد: $F_X(x) = 1 - e^{-\theta x}$ $(x>0)$

الف) احتمال را حساب کنید که یک قطعه از این نوع کالا حداقل ۵ ساعت عمر کند.

ب) چه قدر احتمال دارد این قطعه ۲ ساعت عمر کند؟ * حل. $P(X=2)=0$

ج) متوسط طول عمر این قطعه را حساب کنید.

د) واریانس طول عمر این قطعه را حساب کنید.

حل. الف. روش اول: $f_X(x) = \frac{d}{dx}F_X(x) = \begin{cases} \theta e^{-\theta x} ; & x>0 \\ 0 & o.w. \end{cases}$

$$P(X \ge 5) = \int_5^{\infty} f_X(x)\,dx$$

$$\Rightarrow P(X=5) = \int_5^{\infty} \theta e^{-\theta x}\,dx = -e^{-\theta x}\Big]_5^{\infty} = e^{-5\theta}$$

روش دوم:

$$P(X \ge 5) = 1 - P(X<5) = 1 - F_X(5) = 1-(1-e^{-\theta \times 5}) = e^{-5\theta}$$

حل ج)
$$E(x) = \int_{-\infty}^{\infty} x f(x)\,dx \overset{*}{=} \underbrace{\int_{-\infty}^{0} x f(x)\,dx}_{} + \int_{0}^{\infty} x f(x)\,dx$$

☆ چون $x>0$ داریم $=0$

$$= \int_{0}^{\infty} x\theta e^{-\theta x}\,dx \overset{\text{جزء به جزء}}{=} \frac{1}{\theta}$$

(یا Hop)

نکته ۷) اگر $\theta = 0.001$ ، $E_x = 1000$ شود یعنی عمر متوسط قطعه ۱۰۰۰ ساعت می‌باشد. (x در فرمول برحسب ساعت است)

حل د)
$$E x^2 = \int_{-\infty}^{\infty} x^2 f(x)\,dx \overset{\text{مشابه بالا}}{=} \int_{0}^{\infty} x^2 f(x)\,dx = \int_{0}^{\infty} x^2 \theta e^{-x\theta} = \frac{2}{\theta^2}$$

$$\Rightarrow var(x) = Ex^2 - (E_x)^2 = \frac{2}{\theta^2} - \left(\frac{1}{\theta}\right)^2 = \frac{1}{\theta^2}$$

نکته ۸) اگر $\theta = 0.001$ واریانس 10^6 می‌شود.

مثال. فرض کنید متغیر تصادفی X دارای تابع چگالی احتمال زیر باشد: (دستگاهی محصولاتی تولید می‌کند که قبل از بارگیری بطور صددرصد بازرسی می‌شود، وسیله‌ی اندازه‌گیری محصول طوری است که قادر به خواندن اعداد بین ۱ و $\frac{4}{3}$ نیست. مقدار اندازه‌گیری اندازه‌های به دست آمده دارای چگالی زیر می‌باشند:)

☆ New

$$f_X(x) = \begin{cases} kx^2 & 0 \le x \le 1 \\ \boxed{1} & 1 < x \le \frac{4}{3} \\ 0 & o.w. \end{cases}$$

الف) مطلوبست محاسبه‌ی K.

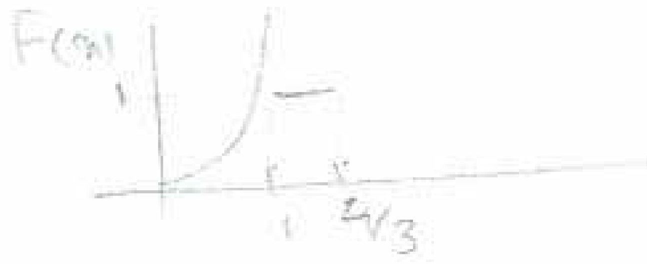
ب) چند درصد از محصولات خارج از فاصله‌ی صفر و یک قرار دارند؟ (احتمال)

ج) امید ریاضی و واریانس این متغیر تصادفی را به دست آورید.

☆ د) تابع توزیع احتمال این متغیر تصادفی را حساب کنید.

SUBJECT:

Year (   ) Month (   ) Date (



حل الف. با توجه به خواص تابع چگالی داریم: $1=\int_{-\infty}^{\infty} f(x)\,dx = 0 + \int_0^1 kx^2\,dx + \int_1^{4/3} 1\,dx$

$$= \frac{k}{3} + \frac{1}{3} = \frac{k+1}{3} \Rightarrow k=2$$

i) $k \geq 0$ and ii) $\int_{-\infty}^{\infty} f(x)\,dx = 1$

☆ جواب مسئله اشتراک این دو شرط است

حل ب. $P(X \notin (0,1)) = 1 - P(0<X<1) = 1 - \int_0^1 2x^2\,dx = 1/3 = 33\%$

روش دوم: $P(X \notin (0,1)) = P(1<X<4/3) = \int_1^{4/3} 1\,dx = 1/3 = 33\%$

$$= P(X \in (1, 4/3))$$

حل ج. $E_X = \int_{-\infty}^{\infty} x f(x)\,dx = \int_{-\infty}^{0} x f(x)\,dx + \int_0^1 + \int_1^{4/3} + \int_{4/3}^{+\infty}$

$$= \int_0^1 x(2x^2)\,dx + \int_1^{4/3} x(1)\,dx = \int_0^1 2x^3\,dx + \int_1^{4/3} x\,dx = \frac{1}{2} + \frac{(4/3)^2-1}{2} = 8/9$$

یعنی متوسط اندازه ی قطعه ۸/۹ است یعنی به طور متوسط اندازه ی هر قطعه ۸/۹ است.

حل د. $F_X(y) = P(X \leq y) \overset{*}{=} \begin{cases} 0 & y<0 \\ \int_0^y 2x^2\,dx = 2y^3/3, & 0<y\leq 1 \\ \int_0^1 2x^2\,dx + \int_1^y 1\,dx = 2/3 + y - 1 & 1<y\leq 4/3 \\ 1 & y \geq 4/3 \end{cases}$

تمرین: متغیر تصادفی X نشان دهنده ی وزن یک کالا است دارای تابع چگالی احتمال زیر می باشد:

$$f(x) = \begin{cases} x-8 & 8 \leq x \leq 9 \\ 10-x & 9 < x \leq 10 \\ 0 & o.w. \end{cases}$$

$E_X = \int_{-\infty}^{\infty} x f(x)\,dx = \int_8^9 (x^2-8x)\,dx + \int_9^{10} (10x - x^2)\,dx$

$= x^3/3\Big|_8^9 - 8x^2/2\Big|_8^9 + 5x^2\Big]_9^{10} - x^3/3\Big|_9^{10} = \dots = 9$

الف) میانگین و واریانس

$Var(x) = E(x^2) - (E_X)^2 = \int_8^9 x^2(x-8)\,dx + \int_9^{10} x^2(10-x)\,dx$

$= x^4/4\Big|_8^9 - 8x^3/3\Big|_8^9 + 10x^3/3\Big|_9^{10} - x^4/4\Big|_9^{10} = \dots \approx 81,16$

☆ پ) تولید کننده ی این کالا آن را به قیمت ثابت ۲۰۰ تومان می فروشد و ضمانت می کند که برگاه

STAEDTLER®

یک مشتری کالایی با وزن ۸٫۲۵ بخرد پول آن را پس دهد. هزینه‌ی تولید کالا بستگی به وزن کالا دارد در واقع

برابر $\frac{35}{100} + \frac{x}{15}$ است. امید ریاضی سود کالا را بدست آورید.

تمرین. تعمیر یک کامپیوتر شخصی بر حسب ساعت یک متغیر تصادفی با تابع چگالی زیر است:

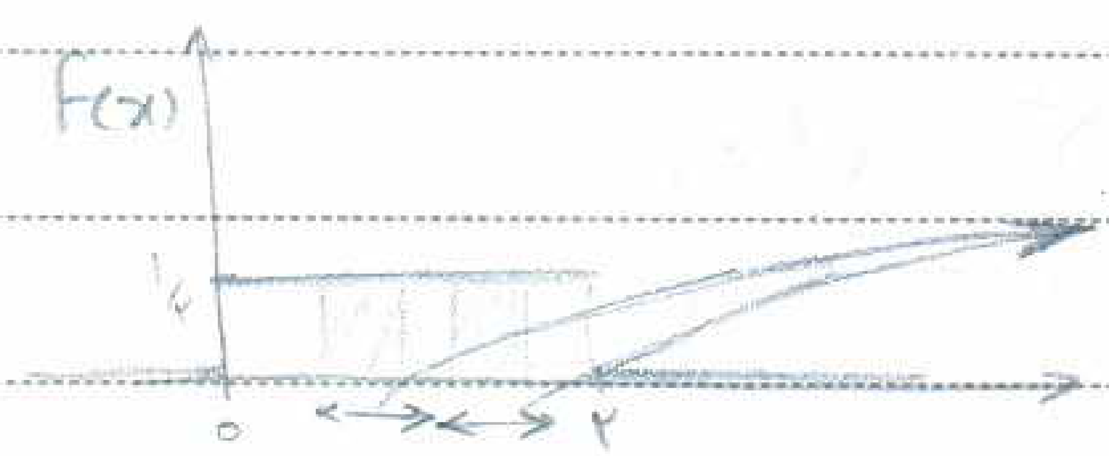
هزینه‌ی تعمیر بستگی به زمان تعمیر دارد و هرگاه زمان $f(x) = \begin{cases} 1/2 & 0<x<2 \\ 0 & o.w. \end{cases}$

x باشد هزینه برابر با $40+30\sqrt{x}$ است. امید ریاضی هزینه‌ی تعمیر کامپیوتر را بدست آورید. (انتگرال)

حل:

$$E(40+30\sqrt{x}) = \int_0^2 \frac{40+30\sqrt{x}}{2} f(x)\,dx =$$

$f(x) = \begin{cases} k & a<x<b \\ \alpha & o.w. \end{cases}$



وقتی چگالی ثابت می‌شود یعنی در تمام بازه‌ها احتمال‌ها برابر است.

+ تذکر (احتمال یعنی سطح زیر منحنی)

☆ اگر در بازه‌های مساوی ... ← متغیر یکنواخت و احتمال‌ها برابر است.

تمرین. تابع چگالی X به صورت زیر است. $f(x) = \begin{cases} a+bx^2 & 0<x\le 1 \\ 0 & o.w. \end{cases}$

اگر $E_X = 3/5$ ، a, b را پیدا کنید.

$$E_X = \int_{-\infty}^{\infty} x f(x)\,dx = \int_0^1 x(a+bx^2)\,dx$$

$$= \int_0^1 ax\,dx + \int_0^1 bx^3\,dx = \frac{ax^2}{2}\Big|_0^1 + \frac{bx^4}{4}\Big|_0^1 = \frac{a}{2} + \frac{b}{4} = \frac{3}{5} \quad (1)$$

طبق خواص تابع چگالی: $\int_{-\infty}^{\infty} f(x)\,dx = 1 \Rightarrow \int_0^1 (a+bx^2)\,dx = ax\Big]_0^1 + \frac{bx^3}{3}\Big]_0^1 = a + \frac{b}{3} = 1 \quad (2)$

$$\Rightarrow \begin{cases} a + b/2 = 6/5 \\ -a - b/3 = -1 \end{cases} \Rightarrow b/6 = 1/5 \Rightarrow b = 6/5 \ , \ a = 3/5$$

$f(x) \ge 0$

تمرین صفحه‌ی قبل. ب) سود کالا = { ... o o o

تابع مولد گشتاور :

فرض کنید X یک متغیر تصادفی و t یک مقدار حقیقی در همسایگی صفر باشد ، در این صورت تابع مولد گشتاور X در نقطه‌ی t به صورت زیر تعریف می‌شود.

$$M_x(t)=E(e^{tx})=\begin{cases}\sum_x e^{tx}P(X=x) & \text{اگر متغیر گسسته باشد :}\\ \int e^{tx}f(x)\,dx & \text{« « پیوسته باشد :}\end{cases}$$

★ بنا به تعریف امید ریاضی $E(X^i)$ را گشتاور مرتبه i ام متغیر تصادفی X می‌گوییم.

توجه شود که $e^a=\sum_{i=0}^{\infty}\frac{a^i}{i!}$ (بسط سری مک لورن تابع $f(x)=e^a$)

در نتیجه

$$e^{tx}=\sum_{i=0}^{\infty}\frac{(tx)^i}{i!}=1+tx+\frac{t^2}{2}x^2+\frac{t^3}{6}x^3+\dots$$

$$\Rightarrow M_x(t)=E(e^{tx})=1+tEx+\frac{t^2}{2}Ex^2+\frac{t^3}{6}Ex^3+\dots$$

$$\Rightarrow \left.\frac{dM_x(t)}{dt}\right|_{t=0}=\left.Ex+tEx^2+\frac{t^2}{2}Ex^3+\dots\right|_{t=0}=Ex$$

$$\left.\frac{d^2M_x(t)}{dt^2}\right|_{t=0}=\left.Ex^2+tEx^3+\dots\right|_{t=0}=Ex^2$$

$$\vdots$$

$$\left.\frac{d^kM_x(t)}{dt^k}\right|_{t=0}=E(x)^k \qquad k=1,2,3,\dots$$

نکته: توجه شود که گشتاورهای یک متغیر تصادفی شاخص‌های مربوط به نحوه‌ی پراکندگی جامعه‌ای را بیان می‌کند که توسط X مشخص می‌شود.

نکته‌ی 2. تابع مولد گشتاور یکتاست. یعنی اگر دو متغیر تصادفی دارای تابع مولد گشتاور یکسانی باشند آنگاه آن دو متغیر تصادفی هم توزیع هستند. * اگر x و y دو متغیر تصادفی گسسته باشند هم توزیعی آن یعنی اینکه: $\forall z ; P(X=z) = P(Y=z)$ $\quad X \overset{d}{=} Y$ هم توزیع

و اگر X و Y پیوسته باشند یعنی اینکه: $\forall z \quad f_X(z) = f_Y(z)$

+ :

$X \to P(X=x) \qquad E(e^{tx}) = \sum_x e^{tx} P(X=x) = \emptyset$

$Y \to P(Y=y) \qquad E(e^{ty}) = \sum_y e^{ty} P(Y=y) = \emptyset$

(شما بعدا) اگر برابر باشند یعنی $P(X=x) = P(Y=y)$ $\Leftarrow X \overset{d}{=} Y$ هم توزیع اند.

مروری بر شاخص‌های جامعه:

توجه شود که هر متغیر تصادفی مانند X یک جامعه را بیان می‌کند. به عبارتی متغیر تصادفی X که بیان عددی یک صفت مورد مطالعه در یک تحقیق آماری است از لحاظ نحوه‌ی پخش شدن روی محور اعداد حقیقی مورد توجه می‌باشد.

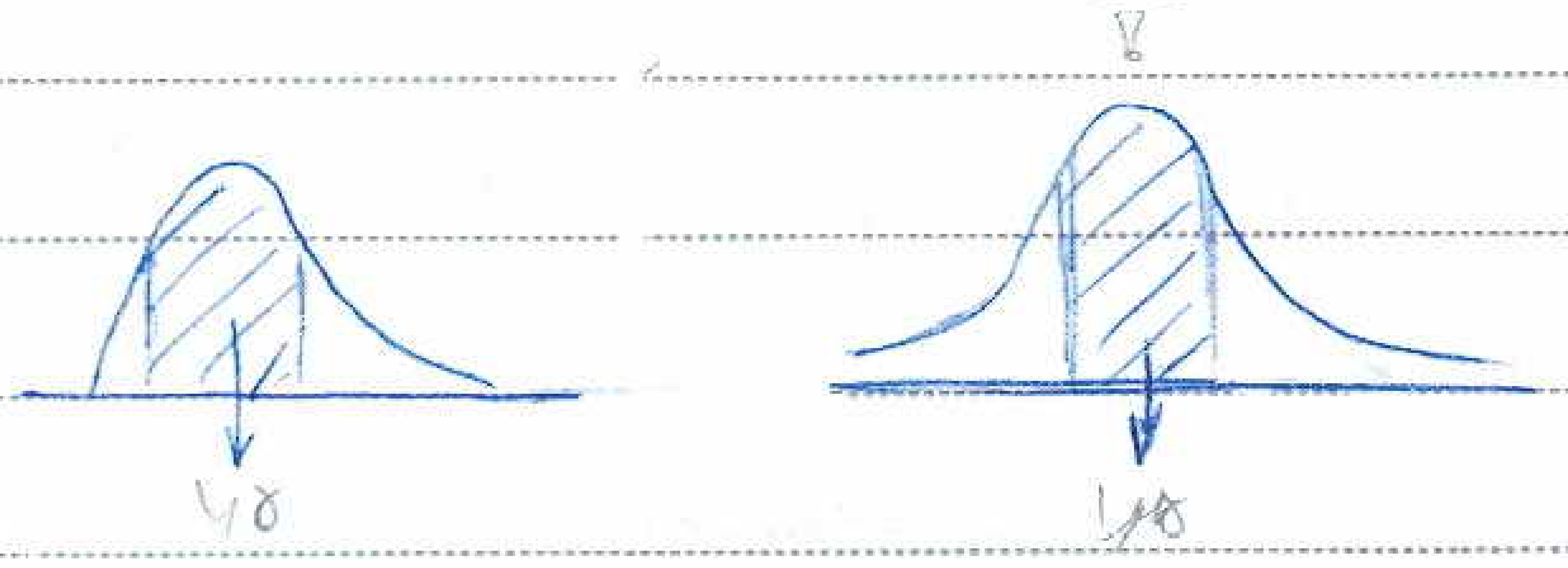
برای یک محقق مهم است که تمرکز مقادیر X کجاست؟ میزان پراکندگی X چه قدر است؟ وضعیت تقارن یا عدم تقارن مقادیر X به لحاظ احتمال (فراوانی‌های نسبی) چگونه است؟ میزان کشیدگی و نرمال بودن
بستگی به احتمالات و فراوانی‌های نسبی دارد

و یا پخ بودن فراوانی‌های نسبی مقادیر X چگونه است؟

* پاسخ کلی سوال‌های فوق را می‌توان براساس شاخص‌هایی که توسط گشتاورهای متغیر تصادفی X ساخته

می شوند، داد.

۱) مهم ترین شاخص تمرکز مقادیر یک متغیر تصادفی مانند X از لحاظ احتمالی، گشتاور مرتبه اول X یا همان امید ریاضی X است. $\mu = E(X) = 1/5$



- یکی دیگر از شاخص های مهم تمرکز میانه است.

بنابه تعریف میانه عددی است که ۵۰٪ جامعه از آن عدد کوچکتر باشند. (جامعه یعنی صفت مورد مطالعه‌ها)

نکته ☆ اگر X یک متغیر تصادفی پیوسته باشد و m میانه ی توزیع X ($f(x)$ ☆ ضابطه ی ریاضی که نشان دهنده ی نحوه ی پخش شدن مقادیر متغیر تصادفی X روی محور اعداد حقیقی به لحاظ احتمالی باشد) باشد، آنگاه m از فرمول زیر محاسبه می گردد:

$$\int_{-\infty}^{m} f(x)\,dx = 1/2$$

☆ سطح زیر منحنی چگالی

(ضابطه ریاضی که از روی هیستوگرام فراوانی (یا درصدی) به دست می آید)

مثال ۱. فرض کنید X دارای تابع چگالی احتمال زیر باشد، میانه ی X را به دست آورید.

$$f(x) = \begin{cases} 3e^{-3x} & x>0 \\ 0 & o.w. \end{cases}$$

$$\int_{-\infty}^{0} f(x)\,dx + \int_{0}^{m} f(x)\,dx =$$

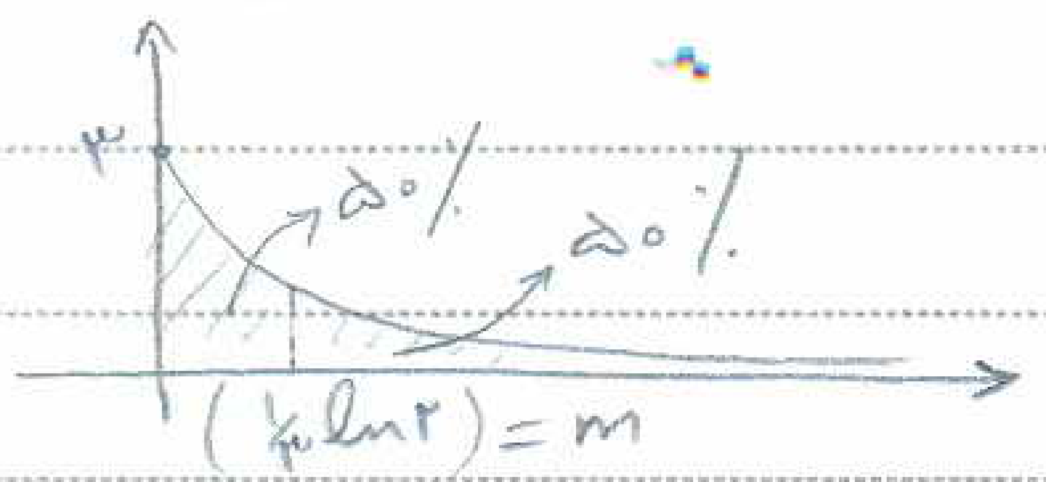
$$= 0 + \int_{0}^{m} 3e^{-3x}\,dx = -e^{-3x}\Big]_{0}^{m} = 1 - e^{-3m}$$

$$\Rightarrow 1/2 = e^{-3m} \Rightarrow -\ln 2 = -3m$$

$$m = 1/3 \ln 2$$

STAEDTLER®

☆ مجموع فراوانی های نسبی = 1 :



- چندک ٤

تعریف - چندک مرتبه $100p$-ام $(0<p<1)$ یک متغیر تصادفی عددی است که $100p\%$ توزیع آن متغیر تصادفی کمتر از آن عدد باشد. آن را با Q_p نشان می‌دهیم. f_p

اگر X یک متغیر تصادفی پیوسته باشد (با تابع چگالی احتمال $f(x)$) آنگاه Q_p از معادلهٔ زیر به دست می‌آید.

$$\int_{-\infty}^{Q_p} f(x)\,dx = p$$

نکته: در بحث چندک‌ها اگر p مقادیر ۰٫۱، ۰٫۲، ...، ۰٫۹ را اختیار کند به آن‌ها دهک می‌گوییم.

اگر p مقادیر ۰٫۰۱، ۰٫۰۲، ...، ۰٫۹۹ را اختیار کند به آن‌ها صدک می‌گوییم.

اگر p مقادیر ۰٫۲۵، ۰٫۵، ۰٫۷۵ را اختیار کند به آن‌ها به ترتیب چارک اول، چارک دوم (میانه) و چارک سوم می‌گوییم.

مثال. فرض کنید X طول عمر یک قطعهٔ الکترونیکی با تابع چگالی احتمال زیر باشد:

$$f(x)=\begin{cases}\theta e^{-\theta x} & x>0 \quad \theta>0\\ 0 & \text{o.w.}\end{cases}$$

مطلوب است محاسبهٔ Q_p.

$$\int_{-\infty}^{Q_p} \underbrace{\theta e^{-\theta x}\,dx}_{f(x)\,dx} = p \Rightarrow \int_0^{Q_p}\theta e^{-\theta} = -e^{-\theta x}\Big|_0^{Q_p} = 1-e^{-\theta Q_p}$$

$\Rightarrow 1-p = e^{-\theta Q_p} \Rightarrow \ln(1-p) = -\theta Q_p \qquad Q_p = -\frac{1}{\theta}\ln(1-p)$

شاخص های پراکندگی :

این شاخص ها میزان دوری و نزدیکی داده ها نسبت به هم را می سنجند. مهم ترین این شاخص ها عبارتند از:

۱) واریانس $\quad Var(X) = E(X-E_X)^2$

۲) انحراف معیار (تبدیل واحد) $\quad \sqrt{Var(X)}$

۳) ضریب تغییرات : فرض کنید X یک متغیر تصادفی باشد. در این صورت ضریب تغییرات X را به صورت زیر تعریف می کنیم

$$C.V(X) = \frac{\sqrt{Var(X)}}{E(X)}$$

* کاربرد: از ضریب تغییرات برای مقایسه ی میزان پراکندگی توزیع دو متغیر تصادفی با واحدهای اندازه گیری متفاوت استفاده می شود.

+ اگر X و Y دو متغیر مفروض و $|CV(X)| < |CV(Y)|$ آنگاه پراکندگی (توزیع) متغیر تصادفی X کمتر است.

شاخص های توزیع :

نحوه ی پخش شدن داده ها روی محور اعداد حقیقی را نشان می دهند و بر دو دسته اند:

الف) شاخص های تقارن : این شاخص ها میزان تقارن یا عدم تقارن داده ها (توزیع متغیر تصادفی) را مشخص می کنند.

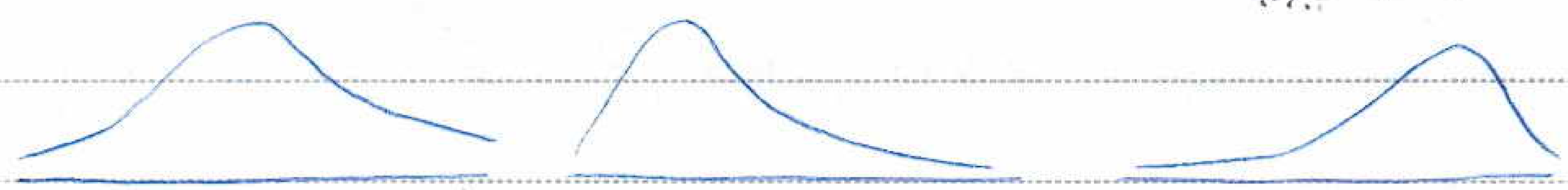
(*) بدون واحد اندازه گیری / تفاوت میانگین ها (E_x و Var ...)

skewness
کجی

مهم‌ترین شاخص تقارن ضریب گشتاوری چولگی است که به صورت زیر ساخته می‌شود.

(Ex^3) ظاهر می‌شود

$$sk = \frac{E(x-Ex)^3}{(\sqrt{Var(x)})^3} = \begin{cases} 0 \rightarrow & \text{توزیع } x \text{ متقارن است.} \\ >0 \rightarrow & \text{توزیع } x \text{ نامتقارن و چاوله به راست است} \\ <0 \rightarrow & \text{توزیع } x \text{ نامتقارن و چاوله به چپ است.} \end{cases}$$



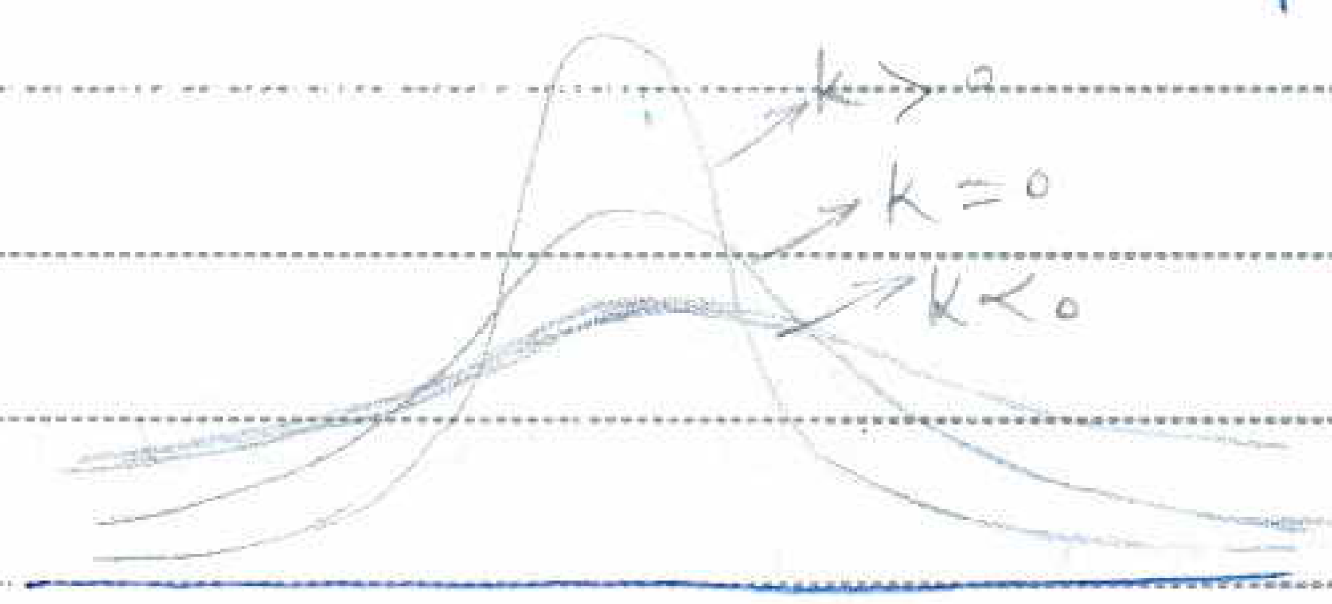
sk = 0   sk > 0   sk < 0

ب) شاخص‌های کشیدگی: این شاخص‌ها میزان کشیدگی داده‌های مختلف را نسبت به کشیدگی مجموعه‌ای از داده‌های خاص که به آن‌ها داده‌های نرمال می‌گوییم می‌سنجند.

مهم‌ترین شاخص کشیدگی ضریب گشتاوری کشیدگی است،

$$k = \frac{E(x-Ex)^4}{(Var(x))^2} - 3$$

$$k = \begin{cases} 0 \rightarrow & \text{داده‌ها نرمال هستند} \\ <0 \rightarrow & \text{داده‌ها پخ‌تر از نرمال هستند.} \\ >0 \rightarrow & \text{داده‌ها کشیده‌تر از نرمال هستند.} \end{cases}$$



- داده‌های نرمال (توزیع نرمال) فرض کنید X یک متغیر تصادفی پیوسته با میانگین μ و انحراف معیار σ باشد (توزیع)

در این صورت می‌گوییم X دارای توزیع نرمال است هرگاه: $(f(x))$

۱- توزیع X متقارن باشد. $sk \approx 0$

۲- تقریباً ۶۷٪ جامعه بین بازه‌ی $(\mu-\sigma , \mu+\sigma)$ قرار گیرد یعنی $P(\mu-\sigma<X<\mu+\sigma) = 0.67$

۳- ۹۵٪ جامعه در بازه‌ی $(\mu-2\sigma , \mu+2\sigma)$ قرار بگیرد.

SUBJECT:
Year ( ) Month ( ) Date ( )

تمرین ...

۴- تقریباً ۹۹٪ جامعه در بازه‌ی $(\mu+3\sigma$ و $\mu-3\sigma)$ قرار گیرد. (۷ ← 4,3,2)

متغیرهای تصادفی توأم:

فرض کنید که فضای نمونه‌ی یک آزمایش تصادفی و X و Y متغیرهای تصادفی باشند که روی S تعریف می‌شوند.

در این بخش رفتار همزمان (توأم) X و Y را مورد بررسی قرار می‌دهیم.

مثال: آزمایش دوبار پرتاب یک سکه را در نظر بگیرید و متغیرهای تصادفی زیر را تعریف کنید.

X : تعداد خط‌ها (در دو پرتاب)   Y : تعداد شیرها در پرتاب اول

تابع جرم احتمال توأم X و Y را به دست آورید.

S = {خخ ، خش ، شخ ، شش}

X : ۲ ۱ ۱ ۰ $\Rightarrow R_X=\{0,1,2\}$ ⇐ X گسسته است

Y : ۰ ۰ ۱ ۱ $R_Y=\{0,1\}$ ⇐ Y گسسته است

نکته ★ برای به دست آوردن تابع جرم احتمال توأم X و Y بایستی به ازای کلیه‌ی مقادیر x و y مقادیر احتمال $P(X=x, Y=y)$ محاسبه شود. به مجموعه‌ی کلیه‌ی این احتمالات "تابع جرم احتمال توأم" X و Y می‌گوییم.

y \ x	0	1	2
0	0	1/4	1/4
1	1/4	1/4	0

← جدول یا تابع جرم احتمال توأم

$\frac{1}{4}+\frac{1}{4}+\frac{1}{4}+\frac{1}{4}=1$ است

$$P(X=x, Y=y)=\begin{cases} \frac{1}{4} & (x,y)\in\{(1,0),(2,0),(0,1),(1,1)\} \\ 0 & \text{o.w.} \end{cases}$$

تعریف . به تابع احتمال $P(X=x, Y=y)$ تابع جرم احتمال توأم متغیرهای تصادفی گسسته ی X و Y

می گوییم هرگاه : الف) $\forall x,y \; ; \; P(X=x, Y=y) \geq 0$

ب) $\sum_y \sum_x P(X=x, Y=y) = 1$

تابع احتمال حاشیه ای :

(در حالت گسسته)

اگر تابع احتمال توأم $P(X=x, Y=y)$ را داشته باشیم ، تابع جرم احتمال متغیر تصادفی گسسته ی X را که به آن

تابع جرم احتمال حاشیه ای X می گوییم با استفاده از قانون احتمال کل به صورت زیر به دست می آوریم :

$$P(X=x) = \sum_y P(X=x, Y=y)$$

✓ $Y=y$ افرازها ... در قانون احتمال کل

و به طور مشابه :

$$P(Y=y) = \sum_x P(X=x, Y=y)$$

- در مثال قبل ، توابع جرم احتمال حاشیه ای X و Y را حساب کنید.

$y \backslash x$	۰	۱	۲	$P(Y=y)$
۰	۰	$1/4$	$1/4$	$1/2$
۱	$1/4$	$1/4$	۰	$1/2$
$P(X=x)$	$1/4$	$1/2$	$1/4$	

$$P(X=x) = \begin{cases} 1/4 & x=0 \\ 1/2 & x=1 \\ 1/4 & x=2 \\ 0 & o.w. \end{cases} = \begin{cases} 1/4 & x=0,2 \\ 1/2 & x=1 \\ 0 & o.w. \end{cases}$$

$$P(Y=y) = \begin{cases} 1/2 & y=0 \\ 1/2 & y=1 \\ 0 & o.w. \end{cases} = \begin{cases} 1/2 & y=0,1 \\ 0 & o.w. \end{cases}$$

SUBJECT:

Year ( ) Month ( ) Date (

تعریف. اگر X و Y دو متغیر تصادفی پیوسته باشند، $f(x,y)$ "تابع چگالی احتمال توأم" X, Y گوییم هرگاه:

الف) $\forall x,y:\ f(x,y) \geq 0$

چون تابع x, y بعدی که عضو کلیه‌گان نباشند

ب) $\iint f(x,y)\,dx\,dy = 1$

تعریف. اگر $f(x,y)$ مفروض باشد، $f(x)$ را که تابع چگالی احتمال حاشیه‌ای X می‌گوییم بطریق زیر محاسبه می‌کنیم:

$$f_X(x) = \int_{R_y} f(x,y)\,dy$$

و به طور مشابه:

$$f_Y(y) = \int_{R_x} f(x,y)\,dx$$

مثال ۲- فرض کنیم X و Y دارای تابع چگالی احتمال زیر باشند. توابع احتمال حاشیه‌ای X, Y را حساب کنید.

الف) $f(x,y) = \begin{cases} 2e^{-x-2y} & x,y>0 \\ 0 & o.w. \end{cases}$

ثابت (در این مثال k را باید پیدا کنید)

ب) $f(x,y) = \begin{cases} k & x,y>0,\ x+y\leq 2 \\ 0 & o.w. \end{cases}$

حل الف.

$$f_X(x) = \int_0^{\infty} f(x,y)\,dy = e^{-x}\int_0^{\infty} 2e^{-2y}\,dy = e^{-x}\times 1$$

$$= e^{-x} \overset{\star}{=} \begin{cases} e^{-x} & x>0 \\ 0 & x\leq 0 \end{cases}$$

$$f_Y(y) \overset{?}{=} \begin{cases} 2e^{-2y} & y>0 \\ 0 & o.w. \end{cases} \qquad = \int_0^{\infty} 2e^{-x-2y}\,dx = -2e^{-2y-x}\Big]_0^{\infty} = +2e^{-2y}$$

SUBJECT: 16

Year ( ) Month ( ) Date ( )

✓ تعمیم موارد ذکر شده برای حالت بیش از دو متغیر ✓

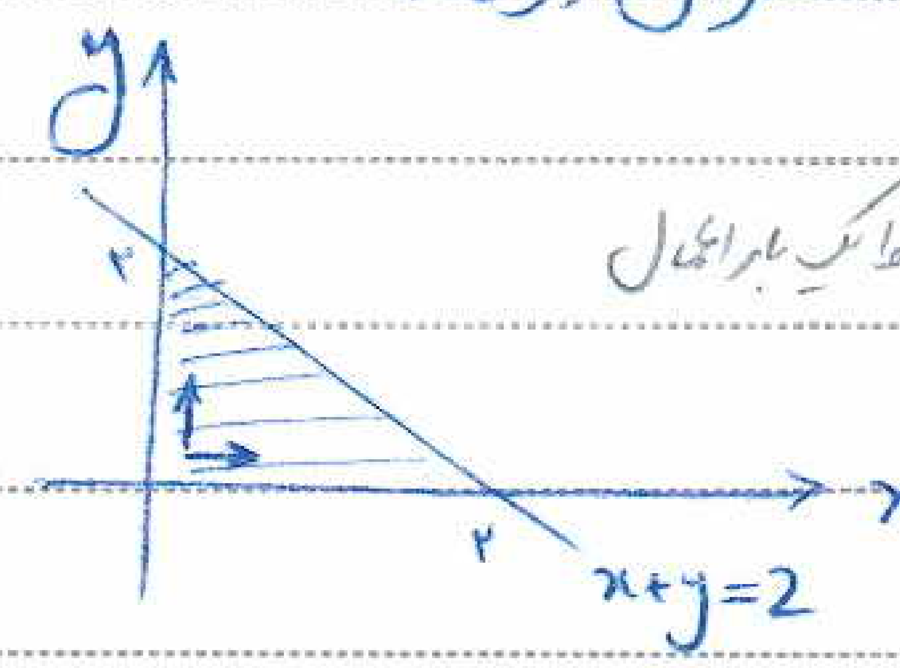
حل ب) * برای پیدا کردن k توجه شود که اولاً $k>0$ ، ثانیاً $\iint f(x,y)\,dx\,dy = 1$.

$$1=\iint f(x,y)\,dx\,dy =$$

* اعمال محدودیت فقط یکبار

$$=\int_0^2\int_{0=x}^{2-y=x} k\,dx\,dy$$

* برای تعیین حدود انتگرال دوگانه :



$$\overset{\text{یا}}{=}\int_0^2 \underbrace{\int_0^{2-x=y} k\,dy}_{\text{داخل}}\,dx$$

۱. * محدودیت مثلاً ۲-y را فقط یکبار اعمال می‌کنیم.

۲. * محدودیت در انتگرال داخلی اعمال می‌شود...

$$= k\int_0^2 (2-x)\,dx = k\left(2x-\tfrac{x^2}{2}\right)\Big]_0^2 = 2k$$

$$\Longrightarrow k=\tfrac{1}{2}$$

$$\Rightarrow f_x(x)=\int_0^{2-x}\underbrace{f(x,y)}_{1/2}\,dy=\tfrac12\int_0^{2-x}dy$$

۳. * برای یک انتگرال مثل اینجا حتماً محدودیت (△ ۲-x) را اعمال می‌کنیم.

$$=\begin{cases}\dfrac{2-x}{2} & 0<x<2\\ 0 & o.w.\end{cases}$$

$$f(y)=\int_0^{2-y}\tfrac12\,dx \overset{*}{=}\begin{cases}\dfrac{2-y}{2} & 0<y<2\\ 0 & o.w.\end{cases}$$

استقلال:

تعریف . دو متغیر تصادفی X و Y را مستقل گوییم اگر و تنها اگر : الف) $R(x,y)=R_x\times R_y$

ب) $F_{X,Y}(x,y)=F_X(x)\times F_Y(y)$

اگر گسسته: $P(X=x, Y=y)$

تذکر $R_{(x,y)} = \{(x,y) \; ; \; f(x,y) > 0\}$

مثال. استقلال X و Y را در مثال ۱ بررسی کنید.

X,Y مستقل نیستند ⇐ ☆ ...

y \ x	0	1	2
0	0	1/4	1/4
1	1/4	1/4	0

$R_x = \{0,1,2\} \qquad R_y = \{0,1\}$

$R_{(x,y)} = \{(1,0),(2,0),(0,1),(1,1)\}$

$R_{(x)} \times R_y = \{(0,0), \ldots, (2,1)\}$

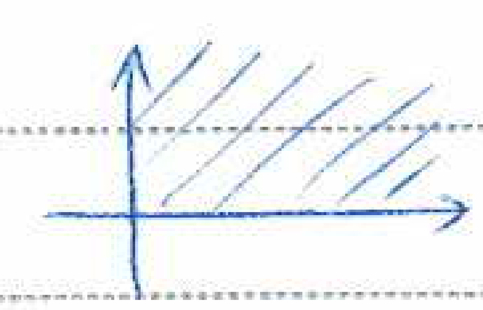
$\Longrightarrow R_{(x,y)} \neq R_x \times R_y \Longrightarrow$ X,Y مستقل نیستند

تذکر تکیه‌گاه یک متغیر تصادفی مجموعه‌ی کلیه‌ی نقاطی است که احتمال در آن نقاط مثبت $(\neq 0)$ باشد، همین تعریف برای تکیه‌گاه توأم دو متغیر تصادفی نیز برقرار است.

تذکر ۲ اگر در یک جدول جرم احتمال توأم حداقل یکی از احتمالات صفر باشد آن دو متغیر تصادفی مستقل از هم نیستند (به هم وابسته‌اند).

مثال. در مثال ۲: الف)

$R_x = (0,\infty) \qquad R_y = (0,\infty)$

$R_x \times R_y = (0,\infty) \times (0,\infty) \xrightarrow{\star}$ 

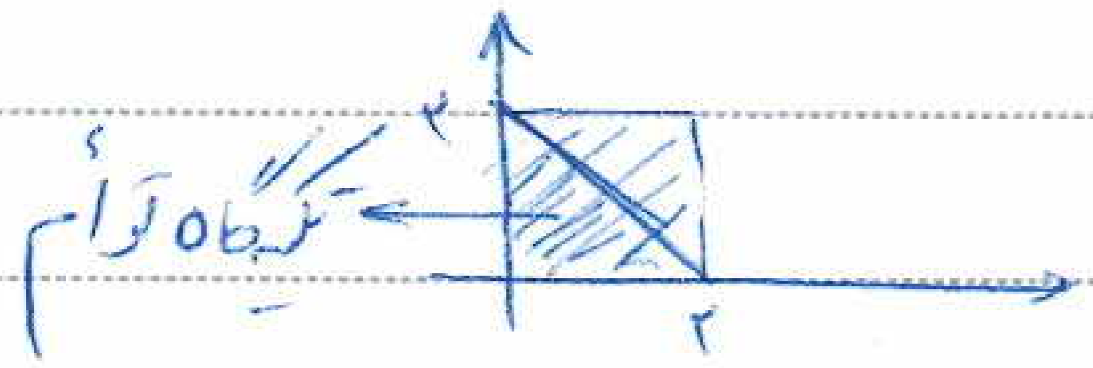
$= R_{(x,y)}$: شرط اول برقرار است

$F_{x,y}(x,y) \overset{\checkmark}{=} F(x) \times F(y) \; \ldots \rightarrow$ شرط دوم هم برقرار است

$\Longrightarrow$ مستقل هستند

در مثال ۲ ب)

$R_x \times R_y \neq R(x,y)$



$\Longrightarrow$ مستقل نیستند

تمرین - در توابع چگالی احتمال توام زیر ، اولاً حاشیه ای ها را حساب کنید.

ثانیاً) استقلال متغیرهای تصادفی مربوطه را بررسی کنید.

$$1)\ f(x,y) = \begin{cases} \frac{12}{5} x(2-x-y) & 0<x<1,\ 0<y<1 \\ 0 & o.w. \end{cases}$$

$$f(x,y) = \begin{cases} x e^{-(x+y)} & x>0,\ y>0 \\ 0 & o.w. \end{cases}$$

$$^{*}f(x,y) = \begin{cases} 2 & 0<x<y,\ 0<y<1 \\ 0 & o.w. \end{cases}$$

$$1)\ f_X(x) = \int_0^1 \frac{12}{5}(2x - x^2 - xy)\,dy = \frac{12}{5}\left(2xy - x^2y - x\frac{y^2}{2}\right)\Big|_0^1 = \frac{12}{5}\left(2x - x^2 - \frac{x}{2}\right) ;\ 0<x<1$$

$$f_Y(y) = \int_0^1 \frac{12}{5}(2x - x^2 - xy)\,dx = \frac{12}{5}\left(x^2 - \frac{x^3}{3} - \frac{x^2}{2}y\right)\Big|_0^1 = \frac{12}{5}(1 - 1/3 - y/2) = \frac{12}{5}(2/3 - y/2) ;\ 0<y<1$$

$$f_X(x)\, f_Y(y) = \frac{144}{25}(2/3 - y/2)\left(2x - x^2 - \frac{x}{2}\right) = \frac{12}{5}x\left(\frac{12}{5} - \frac{8}{5}x - \frac{9}{5}y + \frac{6}{5}xy\right) \neq f(x,y)$$

$$2)\ f_X(x) = \int_0^\infty x e^{-(x+y)}\,dy = -x e^{-(x+y)}\Big|_0^\infty = +x e^{-x} \quad (x>0)$$

$$f_Y(y) = \int_0^\infty x e^{-(x+y)}\,dx = e^{-y}\left[-x e^{-x} - e^{-x}\right]_0^\infty = 1 \times e^{-y} \quad y>0$$

$$\begin{array}{c|c} +x & e^{-x} \\ -1 & -e^{-x} \\ 0 & e^{x} \end{array}$$

$$f_X(x)\, f_Y(y) = f(x,y) \ ,\ f(x,y) \overset{\text{مثبت}}{>} 0 \overset{ask}{\Longrightarrow} R_x R_y = R(x,y) \Longrightarrow \text{مستقل}$$

$$3)\ f_X(x) = \int_0^1 2\,dy = 2y\Big|_0^1 = 2 \quad 0<x<y \qquad f_Y(y) = \int_0^y 2\,dx = 2y ;\ 0<y<1$$

$f_X(x) f_Y(y) \neq f(x,y)$ مستقل نیستند.

$$\lim_{x\to\infty} -x e^{-x} \overset{\infty/\infty}{\underset{Hop}{=}} \lim_{x\to\infty} \frac{-1}{e^x} = 0$$

تابع احتمال شرطی.

فرض کنید X و Y دو متغیر گسسته باشند در این صورت تابع جرم احتمال شرطی X به شرط Y = y را به صورت زیر تعریف می‌کنیم:

$$P(X=x \mid Y=y) \stackrel{*}{=} \frac{P(X=x, Y=y)}{P(Y=y)}$$

اگر X و Y متغیرهای تصادفی پیوسته باشند، تابع چگالی احتمال شرطی X به شرط Y = y را به صورت زیر تعریف می‌کنیم:

$$f_{X|Y}(x \mid y) = \frac{f_{XY}(x, y)}{f_Y(y)}$$

مثال - در آزمایش دوبار پرتاب سکه فرض کنید X تعداد شیرها و Y تعداد خط‌ها در پرتاب اول باشند. تابع جرم احتمال شرطی X به شرط y = 0 را حساب کنید. حل.

y \ x	0	1	2	P(Y=y)
0	0	1/4	1/4	1/2
1	1/4	1/4	0	1/2

$S = \{$خ خ ، خ ش ، ش خ ، ش ش$\}$

X	2	1	1	0
Y	0	1	0	1

$$P(X=x \mid Y=0) = \begin{cases} \frac{0}{1/4} & x=0 \\ \frac{1/4}{1/2} = 1/2 & x=1 \\ \frac{1/4}{1/2} = 1/2 & x=2 \end{cases}$$

$$\Rightarrow \begin{cases} 1/2 & x=1,2 \\ 0 & o.w \end{cases}$$

نکته ۱: تابع جرم (چگالی) احتمال شرطی خود یک تابع جرم (چگالی) احتمال است.

نکته ۲: اگر X و Y مستقل باشند:

$$P(X=x \mid Y=y) \stackrel{\checkmark}{=} P(X=x)$$

SUBJECT:

Year ( ) Month ( ) Date (

تمرین. فرض کنید X و Y دارای تابع چگالی احتمال توأم زیر باشند، چگالی شرطی X به شرط Y را حساب کنید.

$$f(x,y)=\begin{cases} \frac{12}{5}x(2-x-y) & ; 0<x,y<1 \\ 0 & \text{o.w.} \end{cases}$$

$$f(x|y)=\frac{f(X=x,Y=y)}{f(Y=y)}$$

حل. ☆ چک کردن: $\int_0^1\int_0^1 f(x,y)\,dx\,dy \overset{?}{=} 1$

$$f_{X|Y}(x|y)=\frac{6x(2-x-y)}{4-3y}\ , \quad \underbrace{0<x<1}_{☆ \dots}$$

$$f(Y=y)=\int_{R_x} f(x,y)\,dx=\frac{12}{5}\left(1-\frac{1}{3}-\frac{y}{2}\right)=\frac{12}{5}\left(\frac{4-3y}{6}\right)$$

امید ریاضی شرطی.

فرض کنید X و Y دو متغیر تصادفی باشند، در این صورت امید ریاضی شرطی X به Y را از فرمول زیر حساب می‌کنیم.

$$E(X|Y=y)=\begin{cases} \sum_x x\,P(X=x|Y=y) & \text{اگر X و Y گسسته باشند:} \\ \int_{R_x} x\,f_{X|Y}(x|y)\,dx & \text{اگر X و Y پیوسته باشند:} \end{cases}$$

☆ اول حاصل سپس شرطی بعد مقدارگذاری به جای y
الزاماً

مثال - در تمرین قبل مطلوبست محاسبه‌ی $E(X|Y=0/5)$ ؟

$$E(X|Y=0/5)=?=\int_0^1 x f(x|y)\,dx=\frac{1}{4-3y}\int_0^1 6x^2(\overbrace{2-x-y}^{(2-y)-x})\,dx$$

$$=\frac{1}{4-3y}\left\{2(2-y)-\frac{3}{2}\right\}\Big|_{y=0/5}=\frac{3-1/5}{2/5}=0/6$$

* واریانس شرطی.

فرض کنید X و Y دو متغیر تصادفی باشند آنگاه واریانس شرطی X به شرط Y=y را به صورت زیر محاسبه می‌کنیم.

$$Var(X|Y=y)=E\left[(X-E(X|Y=y))^2\right] \qquad E(X-E(X|Y=y))^2$$

$$\overset{?}{=}E[X^2|Y=y]-(E(X|Y=y))^2$$

STAEDTLER®

ص ۱۷۴

$$E(u[X]|y)=\int_{R_x \pm\infty}^{\infty} u(x) f(x|y)\,dx$$

البته گسسته بودن X کافی است.

که در آن:

$$E(x^2|Y=y)=\begin{cases}\sum_x x^2 P(X=x|Y=y) & \text{اگر X و Y گسسته باشند؛}\\ \int x^2 f_{X|Y}(x|y)\,dx & \text{اگر X و Y پیوسته باشند؛}\end{cases}$$

* نکته) اگر X و Y دو متغیر تصادفی مفروض باشند، آنگاه

الف) (امید ریاضی دوگانه) $Ex = E[E(X|Y)]$

ب) $Var(X) = E[var(X|Y)] + Var[E(X|Y)]$

تذکر ۱) امید ریاضی یک متغیر تصادفی مقداری ثابت است. (واریانس یک متغیر تصادفی هم مقداری ثابت است.)

به y بستگی دارد

۲) امید ریاضی و واریانس شرطی X به شرط Y یا مقداری ثابت و یا تابعی از a است. (نسبت به X ثابت است)

اثبات ب)

$$Var[\underbrace{E(X|Y)}_{h(y)}] = E[h(y) - \underbrace{E(h(y))}_{Ex}]^2 = E[h^2(y)] - (E(h(y)))^2 \pm E(x^2)$$

$$= \underbrace{Ex^2 - (Ex)^2}_{Var\,x} - [Ex^2 - E(h^2(y))]$$

$$EE(X|Y) \quad E(E(X|Y))$$

$$\underbrace{E(E(x^2|y)) - E^2(x|y))}_{var(x|y)}$$

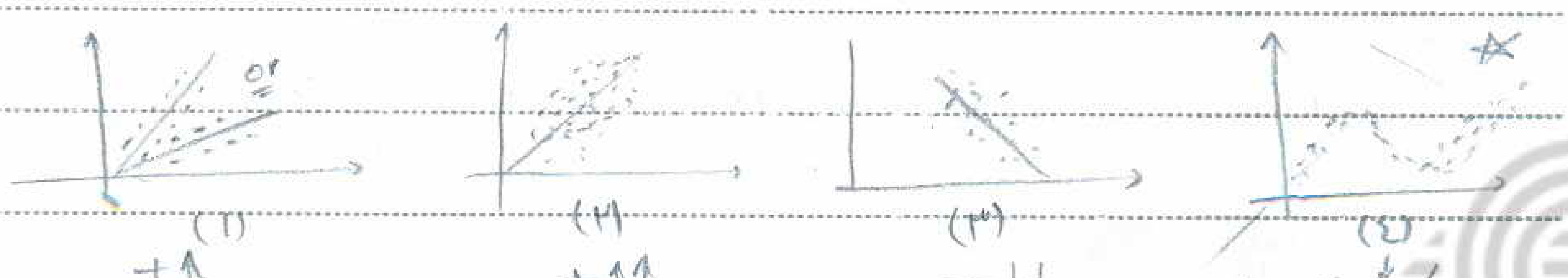
$$[E(x^2) - E(h^2(y))] = E(E(x^2|Y)) - E(E^2(X|Y)) = E(Var(X|Y))$$

۳) * کوواریانس: فرض کنید X و Y دو متغیر تصادفی باشند، کوواریانس X,Y به صورت زیر است:

$$Cov(X,Y) = E\{(X-Ex)(Y-Ey)\}$$

میزان خطی بودن

* شیب مهم نیست فقط (۱) ± بودن شیب مهم است.



(۱) $+\uparrow$ (۲) $+\uparrow\uparrow$ (۳) $-\downarrow\downarrow$ (۴) در این حالت کاری نداریم!

۱/۲۹ ← میان ترم

چهارشنبه هفته اول بعد عید کلاس تشکیل می‌شود. :))

SUBJECT: ضریب همبستگی → { تفاوت واحد ۱) بی دلایل ۲) $cov \in (-\infty, \infty)$ }

Year ( ) Month ( ) Date (

۶) نکته: کوواریانس میزان رابطه خطی بین دو متغیر تصادفی را اندازه می گیرد، اگر کوواریانس عدد مثبتی باشد یعنی یک رابطه خطی مستقیم یا مثبت بین دو متغیر تصادفی وجود دارد.

$cov \pm \equiv$ شیب $\pm$

اگر کوواریانس عددی منفی باشد یعنی یک رابطه ی خطی معکوس یا منفی بین دو متغیر تصادفی وجود دارد

و اگر cov مقداری نزدیک به صفر داشته باشد آنگاه هیچ نوع رابطه ی خطی بین آن دو متغیر تصادفی وجود نخواهد داشت.

نکته: بزرگی و کوچکی عدد کوواریانس بستگی به میزان پراکندگی توده ی نقاط نسبت به هم دارند هرچه توده ی نقاط متراکم تر باشند، اندازه ی واریانس عددی بزرگتر خواهد بود. (قدر مطلق)

$|cov| \equiv$ تراکم

ضریب همبستگی

از این شاخص برای بررسی رابطه ی خطی بین دو متغیر تصادفی استفاده می شود و دقیقاً مانند cov تفسیر می شود اما مشکلات cov را که در زیر آمده است ندارد:

۱) cov دارای واحد اندازه گیری است. (حاصلضرب واحد اندازه گیری X و Y)

۲) cov محدود نیست. یعنی متعلق به $\mathbb{R}$ است.

دلایل فوق باعث می شود نتوان از cov به منظور مقایسه ی شدت و ضعف رابطه ی خطی بین دو زوج

STAEDTLER

متغیر تصادفی استفاده کرد. برای این منظور از ضریب همبستگی با فرمول زیر استفاده می کنیم:

$$\rho(x,y)=\frac{Cov(x,y)}{\sqrt{Var(x)\,Var(y)}}$$

تمرین - نشان دهید: $Cov(x,y) \overset{?}{=} E(xy) - E_x E_y$

$Cov(x,y)=E((x-E_x)(y-E_y)) = E(xy - xE_y - yE_x + E_xE_y)$
$= Exy - E_xE_y + E_xE_y - E_xE_y =$
$= Exy - E_xE_y$ ✓

نکته: اگر x و y دو متغیر تصادفی مستقل باشند آنگاه: $\not\Leftarrow$

۱) $E_{xy} = E_x E_y$

۲) $Cov(x,y) = 0$

۳) $\rho(x,y) = 0$

توجه شود که رابطه ی فوق یک رابطه ی یک طرفه است. (مثال بزنید)

تمرین: ضریب همبستگی و واریانس x و y را برای همه ی مثال های قبل حساب کنید!

- یک معدنچی در معدنی که دارای سه راه خروجی است به دام افتاده است. خروجی اول به تونلی منتهی می شود که پس از ۳ ساعت راهپیمایی نجات می یابد، خروجی دوم وی را پس از ۵ ساعت راهپیمایی به معدن برمی گرداند و خروجی سوم پس از ۷ ساعت او را به معدن برمی گرداند. اگر فرض کنیم معدنچی در تمام حالات با احتمال مساوی یکی از این خروجی ها را انتخاب کند، متوسط زمان نجات وی چه قدر است؟

حل: x زمان نجات فرد
y شماره خروجی انتخاب شده

$E_x \overset{*}{=} E[E(x|y)]$

y	۱	۲	۳
$P(y=y_i)$	۱/۳	۱/۳	۱/۳

SUBJECT: $(E(XEY) = EYEX)$

Year ( ) Month ( ) Date (

$$\Rightarrow E(x) = E[\underbrace{E(x|y)}_{g(y)}] = \sum_{y=1}^{3} g(y)P(y=y) =$$

$$= g(1)\underbrace{P(y=1)}_{1/3} + g(2)\underbrace{P(y=2)}_{1/3} + g(3)\underbrace{P(y=3)}_{1/3}$$

$$= \tfrac{1}{3}\{E(x|y=1) + E(x|y=2) + E(x|y=3)\} = \tfrac{1}{3}\{3 + (5+E_x) + (7+E_x)\}$$

$$= \tfrac{1}{3}\{15 + 2E_x\} \Rightarrow 3E_x = 15 + 2E_x \qquad E_x = 15$$

نکات میانگین و واریانس مجموع متغیرهای تصادفی.

فرض کنید x، y دو متغیر تصادفی و a، b، c مقادیر ثابتی باشند در این صورت:

1) $E(ax+by+c) = aE_x + bE_y + c$

$\overset{?}{=} 0$ $(Var(ax+b) = a^2 Var(x))$

2) $Var(ax+by+c) = a^2 Var(x) + b^2 Var(y) + 2ab\,Cov(x,y)$ (اثبات با نوشتن بر اساس امید ریاضی ... ← پاورقی)

تذکر: اگر $x_1, \dots, x_n$ متغیرهای تصادفی مستقل، $a_1, \dots, a_n$ مقادیر ثابتی باشند آنگاه:

$$Var\left(\sum_{i=1}^{n} a_i x_i\right) = \sum_{i=1}^{n} a_i^2\, Var\, x_i$$

نکته: ← قابل تعمیم و اثبات

$$Cov(a_1x_1 + a_2x_2 + a_3,\ b_1y_1 + b_2y_2 + b_3) = a_1b_1 Cov(x_1,y_1) +$$

$$+ a_1b_2 Cov(x_1,y_2) + a_2b_1 Cov(x_2,y_1) + a_2b_2 Cov(x_2,y_2).$$

پاورقی:

$$var(Z) = E(Z - E_Z)^2 = E(ax+by+\cancel{c} - aE_x - bE_x - \cancel{c})^2 =$$

$$= E(a(x-E_x) + b(y-E_x))^2 = \dots$$

خواص کوواریانس :

۱) $Cov(x,x) = Var(x)$

۲) $Cov(x,a) = 0$
ثابت

۳) $Cov(x,y) = Cov(y,x)$

تمرین - می دانیم اگر x ، y مستقل باشند آنگاه $E(xy) = ExEy$. با یک مثال نشان دهید که عکس حالت فوق برقرار نیست . (جلسه ی بعد)

مثال - فرض کنید x_1 ، x_2 ، ... ، x_n متغیرهای تصادفی مستقل و هم توزیع با تابع چگالی احتمال $f(x)$ و تابع توزیع احتمال $F(x)$ باشند. تابع چگالی احتمال هر یک از ۱) $y = \min(x_1, \dots, x_n)$ ۲) $Z = \max(x_1, \dots, x_n)$ را حساب کنید

حل (۲). ابتدا تابع توزیع z را حساب می کنیم سپس با مشتق گرفتن ...

$$F_Z(z) = P(Z \le z) = P(\max(x_1, \dots, x_n) \le z) =$$

مستقل بودن

$$= P(x_1 \le z, x_2 \le z, \dots, x_n \le z) \overset{*}{=} P(x_1 \le z) P(x_2 \le z) \dots P(x_n \le z)$$

هم توزیعی

$$\overset{*}{=} [P(x_1 \le z)]^n = [F(z)]^n \Longrightarrow f_Z(z) = \frac{d}{dz} F_Z(z) = nf(z)[F(z)]^{n-1}$$

حل (۱) .

$$F_y(y) = P(y \le y) = 1 - P(y \ge y) = 1 - P(\min(x_1, x_2, \dots, x_n) \ge y)$$

$$= 1 - P(x_1 > y, \dots, x_n > y) = 1 - P(x_1 > y) P(x_2 > y) \dots P(x_n > y)$$

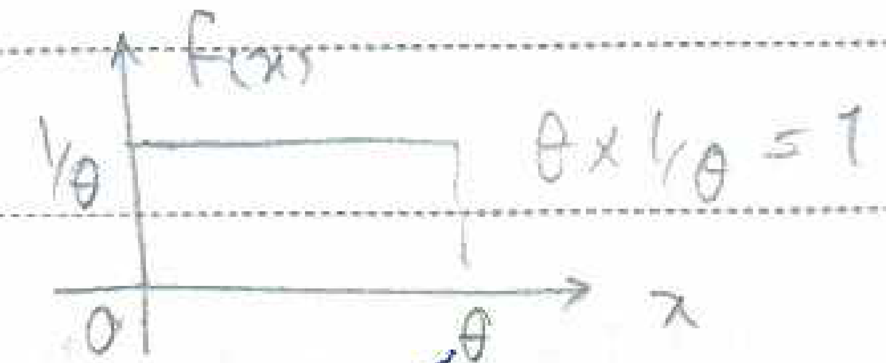
استقلال

$\{X$ هم توزیعی $\overset{?}{=} 1-[P(X_1>y)]^n = 1-[1-F(y)]^n$

$$\Rightarrow f_Y(y) = \frac{d}{dy}F_Y(y) = n f(y)\underbrace{[1-F(y)]}_{\text{تابع بقا (طول عمر)}}^{n-1}$$

مثال ـ فرض کنید $X_1, \dots, X_n$ متغیرهای تصادفی مستقل و هم توزیع با تابع چگالی احتمال زیر باشند که در آن

$$f(x)=\begin{cases} 1/\theta & 0<x<\theta \\ 0 & \text{سایر نقاط} \end{cases}$$



θ مقداری ثابت است.

امید ریاضی (متوسط) ماکزیمم $X_1, \dots, X_n$ را حساب کنید.

حل ـ فرض کنید $Z=\max(X_1,\dots,X_n)$ باید $E(Z)=\int_0^\theta z f_Z(z)\,dz$ را حساب کنیم. برای

این منظور بایستی تابع چگالی Z را حساب کنیم. با استفاده از مثال قبل داریم: $f_Z(z)=n\underbrace{f(z)}_{1/\theta}[F(z)]^{n-1}$

$$F(z)=P(X_1\le z)=\int_0^z \underbrace{f(x)}_{1/\theta}dx=\begin{cases} 0 & z<0 \\ z/\theta & 0<z<\theta \\ 1 & z>\theta \end{cases}$$

$$\Rightarrow f_Z(z)=\frac{n}{\theta}\left(\frac{z}{\theta}\right)^{n-1}=\frac{n z^{n-1}}{\theta^n} \quad (0<z<\theta)$$

$$\Rightarrow E(Z)=\int_0^\theta z\frac{n z^{n-1}}{\theta^n}dz=\frac{n}{\theta^n}\int_0^\theta z^n dz=\frac{n}{\theta^n}\frac{\theta^{n+1}}{n+1}=\frac{n}{n+1}\theta$$

تمرین ـ فرض کنید $X_1, \dots, X_n$ متغیرهای تصادفی مستقل و هم توزیع با تابع چگالی احتمال زیر باشند:

$$f(x)=2e^{-2x} \qquad x>0$$

متوسط مینیمم $X_1 \dots X_n$ را حساب کنید.

حل. X_i: طول عمر قطعه iام در بین n قطعه

$F_Z(z) = P(Z\le z) = P(\min(X_1,\dots,X_n)\le z) =$

$= 1-P(\min(X_1,\dots,X_n)\ge z) = 1-(P(X_i\ge z))^n = 1-[1-P(X_i\le z)]^n = 1-\left(1-\int_0^z f(x)\,dx\right)^n =$

$1-\left(1-\int_0^z 2e^{-2x}\right)^n = 1-\left(1+e^{-2x}\Big]_0^z\right)^n \overset{?}{=} 1-(1+e^{-2z}-1)^n = 1-e^{-2nz}$

$f_Z(z) = \frac{d}{dz}F_Z(z) = 2ne^{-2nz}$ $(z>0) \leftarrow (x>0 \Rightarrow z=\min(X_1,\dots,X_n)>0)$

$E(Z) = \int_0^\infty z f_Z(z)\,dz = \int_0^\infty z\,2ne^{-2nz}\,dz = -ze^{-2nz}\Big]_0^\infty + \int_0^\infty e^{-2nz}\,dz = 0 + \frac{-1}{2n}e^{-2nz}\Big]_0^\infty =$

$= -1/2n$

(چک) $\int_0^\infty f_Z(z)\,dz = -e^{-2nz}\Big]_0^\infty \overset{?}{=} 1$

مثال نقض برای عکس نکته‌ی دو برگه قبل: (نکته بعد ρ)

y \ x	-1	0	1	P(X=x)
-1	1/4	1/3	1/4	2/3
0	0	0	0	0
1	1/4	0	1/4	1/3
P(Y=y)	1/3	1/3	1/3	

$E_X = \frac{-1}{3}+0+\frac{1}{3} = 0$

$E_Y = (-1)\,2/3 + 0\times 0 + 1\times 1/3 = -1/3$

$\Rightarrow Cov(X,Y) = E[(X-E_X)(Y-E_Y)] \overset{*}{=}$

$= (-1)(-1)\cdot 1/4 + 0(-1)\,1/3 + 1(-1)\,1/4 +$

$(-1)\,1\cdot 1/4 + 1\times1\times 1/4 = 0 \rightarrow$ اما دو متغیر مستقل نیستند.

$(=\rho(X,Y))$

همچنین برای $x=-1$, $y=-1$: $E_X E_Y \neq E_{XY}$

$(f(x,y)\neq g(x)h(y))$

نکته

$F(x,y) = P(X\le x, Y\le y) = \int_{-\infty}^{y}\int_{-\infty}^{x} f(s,t)\,ds\,dt$ (ص ۱۱۵)

$E(g(X,Y)) \overset{\text{گسسته}}{=} \sum_x\sum_y g(x,y)f(x,y) \overset{\text{پیوسته}}{=} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f(x,y)\,dx\,dy$

SUBJECT:

Year ( ) Month ( ) Date (

میان ترم: تا آخر اول متغیرهای دوبعدی

(توزیع توابعی از متغیرهای تصادفی)

نمونه — فرض کنید X و Y متغیرهای مستقل و هم توزیع با تابع توزیع چگالی احتمال زیر باشند:

$$f(x) = \begin{cases} e^{-x} & x>0 \\ 0 & o.w. \end{cases}$$

تابع چگالی احتمال $Z = X/Y$ را به دست آورید.

$(V = RI)$ ▲

★ هم توزیع اند یعنی $f(x)$ هر دو یکی است:

$$X \longrightarrow f_X(x) = \begin{cases} e^{-x} & x>0 \\ 0 & o.w. \end{cases}$$

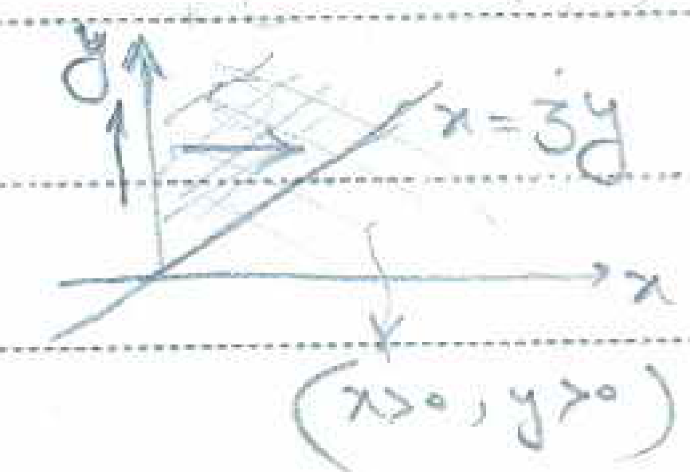
$$Y \longrightarrow f_Y(y) = \begin{cases} e^{-y} & y>0 \\ 0 & o.w. \end{cases}$$

تابع چگالی توام X و Y

$$f_{X,Y}(x,y) \overset{\uparrow}{=} f_X(x)\, f_Y(y) = \begin{cases} e^{-x-y} & x,y>0 \\ 0 & o.w. \end{cases}$$

$$F_Z(\mathfrak{z}) = P(Z \le \mathfrak{z}) = P\left(\underbrace{\frac{X}{Y} \le \mathfrak{z}}\right) \overset{*}{=} \iint\limits_{x/y \le \mathfrak{z}} f(x,y)\,dx\,dy = \text{ادامه}\ldots$$

$$x/y \le \mathfrak{z} \overset{*}{\Longleftrightarrow} x \le \mathfrak{z}y$$



نکته

$$\star\ P(g(x,y) \in B) = \iint\limits_{\underbrace{g(x,y)\in B}} f(x,y)\,dx\,dy$$

(تغییرات x در راستای محور طول)

$$\text{ادامه}\ldots = \int_0^\infty \int_0^{\mathfrak{z}y} e^{-x} e^{-y}\,dx\,dy = \int_0^\infty e^{-y} \underbrace{\left(-e^{-x}\right)\Big|_0^{\mathfrak{z}y}}_{1-e^{-\mathfrak{z}y}} dy =$$

$$= \int_0^\infty \left(e^{-y} - e^{-y(\mathfrak{z}+1)}\right) dy = \left(-e^{-y} + \frac{e^{-y(\mathfrak{z}+1)}}{\mathfrak{z}+1}\right)\Big|_0^\infty = -\left(-1 + \frac{1}{\mathfrak{z}+1}\right) =$$

$$= \frac{\mathfrak{z}}{\mathfrak{z}+1} \quad \mathfrak{z}>0 \longrightarrow \left(\text{چون } x, y>0 \text{ پس } z = \frac{x}{y} > 0\right)$$

بررسی درستی: $\lim\limits_{\mathfrak{z}\to -\infty} F(\mathfrak{z}) = 0 \qquad \lim\limits_{\mathfrak{z}\to +\infty} F(\mathfrak{z}) = 1$ (چهار شرط)

ادامه حل

$$\Longrightarrow f_Z(\mathfrak{z}) = \frac{d}{d\mathfrak{z}} F_Z(\mathfrak{z}) = \frac{\mathfrak{z}+1-\mathfrak{z}}{(\mathfrak{z}+1)^2} = \begin{cases} \dfrac{1}{(\mathfrak{z}+1)^2} & (\mathfrak{z}>0) \\ 0 & o.w. \end{cases}$$

$$\left(\int_0^\infty f_Z(\mathfrak{z}) = 1 \quad : \text{بررسی درستی} + \right)$$

STAEDTLER®

اگر ۱۰.۱ $Z = \frac{X}{X+Y}$ گفته بود: $0 < \mathfrak{z} < 1$

SUBJECT: ۲۳
Year ( ) Month ( ) Date ( )

تک متغیره: میان ترم (تا اول توام در دو متغیره + نامساوی کی‌بی‌شف امروز)

نامساوی مارکوف (Markov):
که در محاسبه‌ی کران ها + برای پیوسته و گسسته برقرار است.

اگر X یک متغیر تصادفی باشد که فقط مقادیر نامنفی را می‌گیرد آنگاه برای هر $a>0$: $P(X\geq a)\leq \frac{EX}{a}$

اثبات برای حالت پیوسته:

$$EX=\int_0^{\infty}xf(x)\,dx=\underbrace{\int_0^{a}xf(x)\,dx}_{\geq 0}+\int_a^{\infty}xf(x)\,dx$$

$$\geq \int_a^{\infty}xf(x)\,dx$$

$$> a\int_a^{\infty}f(x)\,dx$$

چون: $\underset{+}{x}>\underset{+}{a} \Rightarrow \underbrace{\int xf(x)\,dx}_{+}>\underbrace{\int af(x)\,dx}_{+}$ داریم:

$$= aP(X\geq a)$$

$$\Rightarrow EX\geq aP(X\geq a)$$

نامساوی چبیشف (chebychov):

اگر X یک متغیر تصادفی با میانگین μ و واریانس σ^2 باشد، آنگاه برای هر مقدار $k>0$ داریم:

$$P(|X-\mu|\geq k)\leq \frac{\sigma^2}{k^2}$$

به عبارتی:

$$P(|X-\mu|\geq \sigma k)\leq \frac{1}{k^2}$$

اثبات. (با استفاده از نامساوی مارکوف)

$\mu=EX$

$\sigma^2=\operatorname{Var}X=E(X-EX)^2$

بدیهی است که: $(X-\mu)^2$ یک متغیر تصادفی نامنفی است بنابراین براساس نامساوی مارکوف:

$$P((X-\mu)^2\geq k^2)\leq \frac{E(X-\mu)^2}{k^2}=\frac{\sigma^2}{k^2}$$

SUBJECT:

Year ( ) Month ( ) Date (

$$\Rightarrow P(|X-\mu| \geq k) \leq \frac{\sigma^2}{k^2}$$

تمرین - فرض کنید x یک متغیر تصادفی با تابع چگالی احتمال زیر باشد، ضریب چاولگی گشتاوری x را حساب کنید.

$$f(x) = \begin{cases} e^{-x} & x>0 \\ 0 & o.w. \end{cases}$$

★ راه دیگر هم هست ولی این روش را بعنوان یک روش یاد بگیریم؛

$$sk = \frac{E(x-\mu)^3}{\sigma^3} = \frac{E(x^3 - 3x^2\mu + 3x\mu^2 - \mu^3)}{\sigma^3}$$

برای گشتاورهای بالا با انتگرال راحت حساب نمیشه!

$$= \frac{Ex^3 - 3\mu E(x^2) + 3\mu^2 Ex - \mu^3}{\sigma^3 = var^{3/2}}$$

$$\star\ \frac{d}{dt}M_x(t) = Ex \ , \ \frac{d^2}{dt^2}M_x(t) = Ex^2 \ , \ \frac{d^3}{dt^3}M_x(t) = Ex^3$$

$$\sigma^3 = \left(Ex^2 - (Ex)^2\right)^{3/2}$$

- فرض کنید x یک متغیر تصادفی با تابع چگالی احتمال زیر باشد.

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \quad x \in R$$

مطلوبست محاسبه‌ی Ex^2, Ex^3, Ex^4

سعی کنید با تابع مولد گشتاور حساب کنید

چون با f(x) باید چندبار جزء به جزء بزنیم.

$$M_x(t) = e^{t^2/2}$$

...

STAEDTLER®

تمرین ۴ مهم

توزیع‌های خاص احتمالی

الف) توزیع‌های گسسته:

۱- توزیع برنولی. گوییم متغیر تصادفی گسسته‌ی X دارای توزیع برنولی با پارامتر p است هرگاه دارای تابع چگالی
(مشخصه جامعه (توضیحی در مورد آن))
احتمال زیر باشد.

$$P(X=x)=\begin{cases}p^x(1-p)^{1-x} & x=0,1\\ 0 & o.w.\end{cases}$$

مصداق: فرض کنیم در یک آزمایش دو حالتی، موفقیت، شکست، با احتمال موفقیت p ، X تعداد موفقیت‌ها در یک بار انجام آزمایش باشد. $X\sim b(p)$

مثال. فرض کنید ۸۰٪ دانشجویان (رشته مهندسی در) درس آمار قبول می‌شوند، چه قدر احتمال دارد علی در یک کلاس آمار قبول شود؟

حل. ملاحظه می‌شود که آزمایش مورد نظر یک آزمایش دو حالتی است. بنابراین اگر X نشانگر قبول شدن علی باشد آنگاه $X\sim b(0/8)$ دارای توزیع برنولی می‌باشد (با پارامتر ۰/۸)، در نتیجه:

$$P(\text{قبول شدن})=P(X=1)=(0/8)^1(0/2)^0=0/8$$

تمرین ۱. فرض کنید: $X\sim b(p)$ ، الف) نشان دهید: $Ex=p$ ، $Var x=p(1-p)$

ب) تابع مولد گشتاور X را حساب کنید. $M_X(t)=1-p+pe^t$ ← جواب

الف) $Ex=\sum_{x=0,1} x\,P(X=x)=0\times P(x=0)+1\times P(X=1)=1\times p(1-p)^0=p$

$Var x=Ex^2-(Ex)^2=\sum x^2P(X=x)-p^2=p-p^2$

ب) $M_X(t)=\sum_{x=0,1} e^{tx}P(X=x)=P(x=0)+e^tP(X=1)=1-p+e^tp$

$\binom{2n}{n} = \frac{(2n)!}{n!\,n!} = \frac{2n(2n-1)\cdots\times(n+1)}{n!} \times \frac{1}{2^{2n}}$

Subject:

Year. Month. Date. ( )

۲- توزیع دوجمله‌ای:

گوییم متغیر تصادفی گسسته X دارای توزیع دوجمله‌ای با پارامترهای n و p است اگر دارای تابع جرم احتمال زیر باشد:

$$p(X=x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & x=0,1,2,\dots,n \\ 0 & o.w. \end{cases}$$

$\binom{n}{x} = \frac{n!}{x!(n-x)!}$

مصداق: اگر یک آزمایش دوحالتی موفقیت شکست با احتمال موفقیت p را به طور مستقل از هم n بار تکرار کنیم و متغیر تصادفی X را تعداد موفقیت‌ها در n آزمایش در نظر بگیریم (تعریف کنیم) آنگاه X دارای توزیع دوجمله‌ای با پارامترهای n و p است که آن را با نماد زیر نشان می‌دهیم:

$$X \sim B(n,p)$$

تمرین- فرض کنید $X \sim B(n,p)$. الف) نشان دهید: $E_X = np$, $Var_X = np(1-p)$

ب) نشان دهید $M_X(t) = (1-p+e^t p)^n$ ←

ب) $M_X(t) = \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x (1-p)^{n-x}$

$= \sum_{x=0}^{n} \binom{n}{x} \underbrace{(e^t p)^x}_{a} \underbrace{(1-p)^{n-x}}_{b} = (1+e^t p - p)^n$

برای حل ب: بسط دوجمله‌ای نیوتن $(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^i b^{n-i}$

الف)

$E(X) = \sum_{x=0}^{n} x \cdot \binom{n}{x} p^x (1-p)^{n-x} = \sum_{x=1}^{n} \frac{n!}{(x-1)!(n-x)!} p^x (1-p)^{n-x} = np \sum_{x=1}^{n} \binom{n-1}{x-1} p^{x-1} (1-p)^{n-x}$

$\xrightarrow[m=n-1]{y=x-1} = np \sum_{y=0}^{m} \binom{m}{y} p^y (1-p)^{m-y} = np \times 1 = np$

1 = جمع مقادیر تابع دوجمله‌ای با پارامترهای m و p

$Var(X) = E(X^2) - E^2(X) = E(X(X-1)) + E(X) - (EX)^2$

$E(X(X-1)) = \sum_{x=0}^{n} x(x-1) \binom{n}{x} p^x (1-p)^{n-x} = \sum_{x=2}^{n} \frac{n!}{(x-2)!(n-x)!} p^x (1-p)^{n-x} = n(n-1)p^2 \sum_{x=2}^{n} \binom{n-2}{x-2} p^{x-2} (1-p)^{n-x}$

$\xrightarrow[m=n-2]{y=x-2} = n(n-1)p^2 \sum_{y=0}^{m} \binom{m}{y} p^y (1-p)^{m-y} = n(n-1)p^2$

$\Rightarrow Var(X) = n(n-1)p^2 + np - n^2p^2 = np(1-p)$ ✓

Subject: ۲۵

Year. Month. Date. ( )

اثبات قضیه ٤

قضیه) اگر $X_1, \dots, X_n$ متغیرهای مستقل و هم توزیع از توزیع $b(p)$ باشند آنگاه $X=\sum_{i=1}^{n} X_i \sim B(n,p)$

اثبات: می‌دانیم اگر $X \sim B(n,p)$ آنگاه $M_X(t)=(1-p+pe^t)^n$ حال اگر

$M_{\sum_{i=1}^n X_i}(t)=(1-p+pe^t)^n$ آنگاه بنا به منحصر به فرد بودن تابع مولد گشتاور می‌توانیم نتیجه بگیریم که $\sum_{i=1}^{n} X_i \sim B(n,p)$:

اما:

$$M_{\sum_{i=1}^n X_i}(t)=E\left[e^{t\sum_{i=1}^n X_i}\right]=E\left[\prod_{i=1}^n e^{tX_i}\right]\overset{\{X_i\}\text{ استقلال}}{=}\prod_{i=1}^n \underbrace{E(e^{tX_i})}_{M_{X_i}(t)}$$

$$\overset{\{X_i\}\text{ هم توزیعی}}{=}\left(M_{X_1}(t)\right)^n \overset{\text{بنا به تمرین ۱}}{=}(1-p+pe^t)^n$$

نکته) اگر $X_1, \dots, X_n$ متغیرهای تصادفی مستقل باشند و g یک تابع حقیقی باشد آنگاه:

$g(X_1), \dots, g(X_n)$ نیز متغیرهای تصادفی مستقل هستند.

مثال. اگر شخصی یک آزمون ۴ گزینه‌ای را که شامل ۱۰۰ سوال است به تصادف پاسخ دهد:

الف) چقدر احتمال دارد ۵۰٪ سوالات را صحیح پاسخ دهد؟

ب) چقدر احتمال دارد حداقل به دو سوال صحیح پاسخ دهد؟

ج) انتظار دارید چند پاسخ صحیح بدهد؟

حل.

$X \sim B(100, \frac{1}{4})$: تعداد پاسخ‌های صحیح.

الف) $50\% \times 100 = 50$ $\quad P(X=50)=\binom{100}{50}\left(\frac{1}{4}\right)^{50}\times\left(\frac{3}{4}\right)^{50}$

Subject:

Year. Month. Date. ( )

ب) $P(X\geq 2) = P(X=2$ یا $X=3$ یا ... یا $X=100) \overset{\text{دو به دو ناسازگارند}}{=} \sum_{x=2}^{100} P(X=x)$

or $P(X\geq 2) = 1-P(X<2) = 1-P(X=0$ یا $X=1) \overset{\text{دو به دو ناسازگار}}{=} 1-P(X=0)-P(X=1)$

$= 1-\binom{100}{0}(1/4)^{0}(3/4)^{100} - \binom{100}{1}(1/4)^{1}(3/4)^{99} = 1-(3/4)^{100} - 25\times(3/4)^{99}$

ج) $E(X) = np = 100\times\frac{1}{4} = 25$

۳- توزیع هندسی (Geometric)

گوییم متغیر تصادفی گسسته ی X دارای توزیع هندسی با پارامتر p است هرگاه دارای تابع جرم احتمال زیر باشد:

$$P(X=x) = \begin{cases} p(1-p)^{x-1} & x=1,2,\dots \\ 0 & o.w. \end{cases}$$

مصداق: در یک آزمایش دو حالتی موفقیت شکست، با احتمال موفقیت p اگر متغیر تصادفی X را تعداد آزمایش های لازم تا حصول اولین موفقیت تعریف کنیم آنگاه X دارای توزیع هندسی با پارامتر p است که آن را با نماد زیر نشان می دهیم:

$X \sim G(p)$

$\lim_{n\to\infty} q+q^2+\dots = \frac{1-p}{1-1+p} = 1/p-1$

مثال: اگر $X\sim G(p)$ ثابت کنید: $Ex = 1/p$, $\text{Var}x = \frac{1-p}{p^2}$

تمرین ۲: تابع مولد گشتاور توزیع $G(p)$ را حساب کنید. راهنمایی: $(1-p)=q$

$\sum_{x=1}^{\infty} (qe^t)^x$ $\sum_{x=1}^{\infty} pq^{x-1} = 1$ $1+2q+3q^2+\dots = \frac{1}{(1-q)^2}$

برای Var: $1+4q+9q^2+\dots = \frac{1+q}{(1-q)^3}$

$1+q+q^2+\dots = \frac{1}{1-q}$

Subject: ۲۶

Year. Month. Date. ( )

$$\frac{p}{q}\sum_{x=1}^{\infty} q^{x}=1 \Rightarrow \sum_{x=1}^{\infty} q^{x}=\frac{q}{p}=\frac{1-p}{p}=\frac{q}{1-q}$$

$$\sum_{x=1}^{\infty} a^{x}=\frac{a}{1-a}$$

یا سری هندسی ...

تابع مولد گشتاور: $M_{t}(x)=\sum_{1}^{\infty} e^{tx}\, p(1-p)^{x-1}=\sum_{1}^{\infty} \underbrace{(e^{t}q)^{x}}_{\substack{a+a^2+\cdots \\ a(1+a+\cdots)}}\frac{p}{q}=\frac{e^{t}q}{1-e^{t}q}\times\frac{p}{q}=\frac{pe^{t}}{1-e^{t}q}$ ✓

مثال: تیراندازی با احتمال ۰٫۷ به هدف می‌زند چقدر احتمال دارد اولین تیری که به هدف می‌خورد پنجمین شلیک وی باشد؟ انتظار داریم در 20 بار شلیک چند بار هدف مورد اصابت قرار گیرد؟ (فرض براینست که شلیک‌ها مستقل از هم‌اند)

حل. $X \sim G(0.7)$: تعداد شلیک‌های لازم

$P(X=5)=pq^{4}=0.7\times(0.3)^{4}$

A_i: پیشامد اینکه شلیک iام به هدف بخورد، $i=1,2,\dots,5$

$P(A_1^c\cap A_2^c\cap A_3^c\cap A_4^c\cap A_5)=$

$=\underbrace{P(A_1^c)P(A_2^c)P(A_3^c)P(A_4^c)}_{(0.3)^4}\underbrace{P(A_5)}_{0.7}$

$E(x)=\frac{1}{p}=\frac{1}{0.7}=\frac{10}{7}\simeq 1.43 \qquad \frac{20}{1.43}=\frac{20}{\frac{1}{0.7}}=14$

دوم ✓

۴- توزیع پاسکال (دوجمله‌ای منفی)

گوئیم متغیر تصادفی X دارای توزیع پاسکال یا دوجمله‌ای منفی با پارامترهای r و p است هرگاه دارای تابع جرم احتمال زیر باشد:

$$P(X=x)=\begin{cases}\binom{x-1}{r-1}p^{r}q^{x-r} & x=r,r+1,\dots \\ 0 & o.w.\end{cases}$$

مصداق: اگر در یک دنباله از آزمایش‌های دوحالتی موفقیت شکست، با احتمال موفقیت p که به طور مستقل

$M_{\sum_{i=1}^{r} X_i}(t) = E(e^{t\sum X_i}) = E(\prod_{i=1}^{r} e^{tX_i}) = \prod_{i=1}^{r} E(e^{tX_i}) = \left(\frac{pe^t}{1-qe^t}\right)^r$

ازهم انجام می شوند. متغیر تصادفی X را برابر تعداد آزمایش های لازم تا حصول r مین موفقیت در نظر بگیریم، آنگاه X دارای تابع توزیع دوجمله ای منفی (دوجمله ای) با پارامترهای r و p است که آن را با نماد زیر نشان می دهیم:

$$X \sim NB(r,p)$$

نکته) اگر $X \sim NB(r,p)$ آنگاه $E_X = \frac{r}{p}$ ، $Var(X) = \frac{rq}{p^2}$. $(q=1-p)$

تمرین (۱). تابع مولد گشتاور $NB(r,p)$ را حساب کنید و با استفاده از آن نکته ی قبل را ثابت کنید.

$\left(\frac{pe^t}{1-qe^t}\right)^r$

تمرین (۲). اگر $X_1,\dots,X_r \overset{iid}{\sim} G(p)$ آنگاه: $Y=\sum_{i=1}^{r} X_i \sim NB(r,p)$

مستقل و هم توزیع با توزیع $G(p)$

مثال. ظرفی محتوی ۵ مهره ی سفید، ۳ مهره ی آبی و ۲ مهره ی قرمز است. در یک بازی بین پدر و پسر، اگر پسر بتواند حداکثر در ۵ بار مهره برداری با جایگذاری (مستقل) به ۳ مهره ی قرمز برسد برنده است. احتمال برنده بودن وی را حساب کنید.

حل. ملاحظه می شود که آزمایش دوحالتی با احتمال موفقیت $p = 2/10$ است و از آنجایی که قرار است به ۳ مهره قرمز دست یابیم، بنابراین: X = تعداد آزمایش های لازم : $X \sim NB(3, 0.2)$

$$\Rightarrow P(\text{برد پسر}) = P(X \le 5) = P(X=3 \text{ یا } X=4 \text{ یا } X=5)$$

$$P(X=x) = \begin{cases} \binom{x-1}{2}(0.2)^3(0.8)^{x-3} & x=3,4,\dots \\ 0 & o.w. \end{cases}$$

$P(X=n) = \binom{2n}{n}\left(\frac{1}{2}\right)^{2n} \xrightarrow[n\to\infty]{?} 1/2$

$$= P(X=3) + P(X=4) + P(X=5) = (0.2)^3\left[1 + \underbrace{3(0.8)}_{2.4} + \underbrace{6(0.8)^2}_{3.84}\right] =$$

$$= 0.008 \times 7.24 = 0.05792$$

ب) انتظار داریم ۳ مهره ی قرمز در چند بار تکرار آزمایش به دست آید.

$$E_X = \frac{r}{p} = \frac{3}{0.2} = 15$$

+ مثال کتاب

۵- توزیع پواسون.

گوییم متغیر تصادفی X دارای توزیع پواسون با پارامتر λ است هرگاه دارای تابع جرم احتمال زیر باشد.

$$P(X=x) = \begin{cases} \dfrac{e^{-\lambda}\lambda^x}{x!} & x = 0, 1, 2, \dots \\ 0 & \text{o.w.} \end{cases}$$

مصداق: فرض کنید متوسط تعداد اتفاقاتی که در یک بازه ی زمانی مشخص رخ می دهد، برابر λ باشد. اگر متغیر تصادفی X را تعداد اتفاقاتی در نظر بگیریم که در بازه ای با همان طول رخ می دهد، آنگاه X دارای توزیع پواسون با پارامتر λ است که آن را با نماد روبرو نشان می دهیم: $X \sim P(\lambda)$

تمرین ۱. تابع مولد گشتاور $X \sim P(\lambda)$ را محاسبه کنید.

تمرین ۲. نشان دهید اگر $X \sim P(\lambda)$ آنگاه $E(X) = Var(X) = \lambda$. (تحویل فردی)

تذکر: توجه شود که توزیع پواسون تنها توزیعی است که امید ریاضی و واریانس آن با هم برابرند.

Subject:
Year. Month. Date. ( )

مثال. فرض کنید میانگین و واریانس یک متغیر تصادفی برابر ۵ است. احتمال اینکه مقدار آن متغیر تصادفی برابر ۳ مشاهده شود، چه قدر است؟

Ex

حل. $Ex = Var\,x = 5 \Rightarrow x \sim P(5) \qquad P(x=3) = \frac{e^{-5}5^3}{3!}$

مثال ۲. فرض کنید متوسط تعداد تلفنهایی که در هر ۱۰ دقیقه به یک شرکت می شود برابر ۵ تلفن باشد، اگر منشی این شرکت بین ساعت ۱۰ تا ۱۰:۱۰ محل کار خود را ترک کند، چه قدر احتمال دارد در این مدت ۳ تلفن بی پاسخ داشته باشیم؟

(λ)

x: تعداد تلفنها در هر ۱۰ دقیقه $\sim P(5) \Rightarrow P(x=3) = \frac{e^{-5}5^3}{3!}$

فرآیند پواسون (تعمیم توزیع پواسون)

اگر λ (متوسط تعداد اتفاقات) در یک واحد زمانی محاسبه شده باشد، برای مطالعه ی تعداد اتفاقات در t واحد زمانی، $N(t)$، توجه شود که بنا به خواص یک فرآیند پواسون، تابع جرم احتمال $N(t)$ به صورت زیر است.

$t>0$

متغیر X!

$$P[N(t)=x] = \begin{cases} \frac{e^{-\lambda t}(\lambda t)^x}{x!} & x=0,1,2,\dots \\ 0 & o.w. \end{cases}$$

$(\lambda \longrightarrow \lambda t)$

درنتیجه: $E(N(t)) = Var(N(t)) = \lambda t.$

تمرین. فرض کنید تعداد اتوبوس هایی که در هر ۵ دقیقه به یک ایستگاه اتوبوس وارد می شود، به طور متوسط ۶ اتوبوس باشد. چه قدر احتمال دارد: الف) بین ساعت ۸ تا ۸:۱۰، ۱۰ اتوبوس وارد ایستگاه شود؟

ب) بین ساعت ۱۲ تا ۱۲:۱۵، ۱۲ اتوبوس وارد ایستگاه شود؟

ج) بین ساعت ۱:۵۵ تا ۲، ۲ اتوبوس وارد ایستگاه شود؟

د) در هر حالت انتظار دارید چند اتوبوس وارد ایستگاه شود؟

حل. ۵ دقیقه = 1 واحد زمانی الف) $t=2$ ب) $t=\frac{17}{5}=3{,}4$

ج) $t=1$

(me)

الف) $P(N(t)=10)=\frac{e^{-2\times 6}(2\times 6)^{10}}{10!}$ $Ex=\lambda t=6\times 2$ ج) $P(X=2)=\frac{e^{-6}6^2}{2!}$ ...

ب) $P(N(t)=12)=\left(e^{-3{,}4\times 6}(3{,}4\times 6)^{12}\right)\big/ 12!$ $Ex=6\times 3{,}4$

تمرین - فرض کنید $X_1,\dots,X_n \overset{iid}{\sim} P(\theta)$ ثابت کنید $Y=\sum_{i=1}^{n}X_i\sim P(n\theta)$.

$$M_Y(t)=M_{\sum_{i=1}^n X_i}(t)=E\left(e^{t\sum_{i=1}^n X_i}\right)=E\left(\prod_{i=1}^n e^{tX_i}\right)=\prod_{i=1}^n \underbrace{E(e^{tX_i})}=\left(e^{\lambda(e^t-1)}\right)^n \rightarrow \text{ادامه...}$$

$P(\lambda)$

$$M_X(t)=\sum_{x=0}^{n}e^{tx}\frac{e^{-\lambda}\lambda^x}{x!}=e^{-\lambda}\underbrace{\sum_{x=0}^{n}\frac{(\lambda e^t)^x}{x!}}_{\text{سری مکلارن با } e^z,\ z=\lambda e^t}=e^{-\lambda}e^{\lambda e^t}=e^{\lambda(e^t-1)}$$

★ تابع مولد گشتاور توزیع پواسون: $M_X(t)=e^{\lambda(e^t-1)}$ ص ۲۰۵ فروند

$\rightarrow$ ادامه... $t=n$ (تعمیم) اینجا $\Rightarrow \left(=e^{\lambda t(e^t-1)}\ \checkmark\right)$

$(P\in n\theta)$: $$M_X(t)=\sum_{x=0}^{n}e^{tx}\frac{e^{-\lambda t}(\lambda t)^x}{x!}=e^{-\lambda t}\sum_{x=0}^{\infty}\frac{(e^t\lambda t)^x}{x!}\overset{\text{مکلارن}}{=}e^{-\lambda t}e^{e^t\lambda t}=e^{\lambda t(e^t-1)}$$

اعمال t در فرمول
تعمیم توزیع پواسون

$$\frac{d^2M}{dt^2} \stackrel{*}{=} \frac{d}{dt}\left(\frac{\theta}{(\theta-t)^2}\right) = \frac{2\theta}{(\theta-t)^3} \xrightarrow[\text{در همسایگی صفر است}]{\text{مقدار حقیقی } t \text{ در}} = \frac{2}{\theta^2} \Rightarrow Var(x) = \frac{2}{\theta^2} - \left(\frac{1}{\theta}\right)^2 = \frac{1}{\theta^2}$$ تمرین (ادامه):

Subject:

Year. Month. Date. ( )

ب) توزیع‌های احتمالی پیوسته

۱- توزیع نمایی. گوئیم متغیر تصادفی پیوسته‌ی x دارای توزیع نمایی با پارامتر θ است هرگاه دارای تابع چگالی احتمال زیر باشد

$$f(x) = \begin{cases} \theta e^{-\theta x} & x > 0 \\ 0 & o.w. \end{cases} \quad (\alpha = 1) \qquad \text{(Exponential)}$$

مصداق: فرض کنید متوسط زمان لازم تا رخداد احتمالی (اتفاقی) خاص برابر $1/\theta$ باشد، اگر متغیر تصادفی X را زمان لازم تا رسیدن به اولین حادثه یا اتفاق از همان نوع در نظر بگیریم آن‌گاه X دارای توزیع نمایی با پارامتر θ است که آن را با نماد $X \sim Exp(\theta)$ نشان می‌دهیم.

تمرین. فرض کنید $X \sim Exp(\theta)$؛ الف) تابع مولد گشتاور X را حساب کنید.

$$M_X(t) = \int_0^\infty e^{tx}\theta e^{-\theta x}\,dx = \theta\int_0^\infty e^{(t-\theta)x}\,dx = \frac{\theta}{t-\theta}e^{(t-\theta)x}\Big|_0^\infty \xrightarrow[\text{در همسایگی صفر}]{\text{مقدار حقیقی } t} = 0 - \frac{\theta}{t-\theta}e^0 = \frac{\theta}{\theta-t}$$

ب) نشان دهید: $Ex \stackrel{*}{=} 1/\theta$ ، $Var(x) = \frac{1}{\theta^2}$

$$\Rightarrow Ex = \frac{dM}{dt} = \frac{\theta}{(\theta-t)^2} \xrightarrow[\text{در همسایگی صفر}]{\text{مقدار حقیقی } t} = \theta/\theta^2 = 1/\theta \checkmark$$

(راه ۲) با جزء به جزء و تعریف

مثال. فرض کنید متوسط طول عمر لامپهای تولیدی یک کارخانه ۱۰۰۰ ساعت باشد. اگر شخصی لامپی از تولیدات این کارخانه بخرد؛ الف) چه قدر احتمال دارد که این لامپ بیشتر از ۱۲۰۰ ساعت عمر کند؟

ب) چه قدر احتمال دارد این لامپ بین ۸۰۰ تا ۱۵۰۰ ساعت عمر کند؟

ج) انتظار داریم این لامپ چه قدر عمر کند؟

حل.
$$1/\theta = 1000 \Rightarrow \theta = 0/001 \qquad X \sim Exp(0/001)$$

$$f(x) = \begin{cases} 0/001\, e^{-0/001x} & x \geq 0 \\ 0 & o.w. \to x < 0 \end{cases}$$

★ توزیع گاما: $M_x(t) = \left(\frac{\theta}{\theta - t}\right)^{\alpha}$, $E_x = \frac{\alpha}{\theta}$, $Var x = \frac{\alpha}{\theta^2}$

الف) $P(X > 1200) = \int_{1200}^{\infty} 0.001 e^{-0.001x} dx = -e^{-0.001x}\Big]_{1200}^{\infty} = e^{-1.2} < 1$

ب) $P(800 < X < 1500) = \int_{800}^{1500} 0.001 e^{-0.001x} dx = -e^{-1.5} + e^{-0.8}$

ج) $E_x = \frac{1}{\theta} = 1000$

۲- توزیع گاما: گوئیم متغیر تصادفی پیوسته ی X دارای توزیع گاما با پارامترهای α و θ است هرگاه دارای تابع چگالی احتمال روبرو باشد:

$$f(x) = \begin{cases} \frac{\theta^{\alpha} x^{\alpha-1} e^{-\theta x}}{\Gamma(\alpha)} & x > 0, \quad \alpha, \theta > 0 \\ 0 & o.w. \end{cases}$$

که در آن $\Gamma(\alpha)$ «تابع گاما» است که به صورت زیر تعریف می شود: $\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx$

اگر $\alpha \in \mathbb{N}$ آنگاه ثابت می شود که: $\Gamma(\alpha) = (\alpha - 1)!$

مصداق: توزیع گاما حالت تعمیم یافته ی توزیع نمایی است که برای متغیرهایی که نشان دهنده ی طول عمر باشند می توانند استفاده شوند، و حالت خاص که $\alpha \in \mathbb{N}$، مدت زمان لازم تا حصول α امین اتفاق در یک آزمایش طول عمر که متوسط زمان لازم برای هر اتفاق برابر $\frac{1}{\theta}$ است دارای توزیع گاما با پارامترهای α و θ می باشد. نماد: $X \sim \Gamma(\alpha, \theta)$

مثال. فرض کنید متوسط زمان لازم تا رسیدن یک اتوبوس به یک ایستگاه دو دقیقه باشد، چه قدر احتمال دارد حداکثر ۸ دقیقه طول بکشد تا ۵ اتوبوس وارد ایستگاه شوند؟

Subject:
Year. Month. Date. ( )

حل. X: مدت زمان لازم تا رسیدن پنجمین اتوبوس. $X \sim \Gamma(5, 1/2)$ $\quad \frac{1}{\theta} = 2 \Rightarrow \theta = 1/2$

$$P(X \le 8) = \int_0^8 f(x)\,dx = \int_0^8 \frac{x^4 e^{-x/2}}{2^5 \times 4!}\,dx \overset{\text{ج.ج}}{=} \dots$$

$$\begin{array}{c|c} + \; x^4 & e^{-x/2} \\ \hline - \; 4x^3 & -2e^{-x/2} \\ + \; 12x^2 & 4e^{-x/2} \\ - \; 24x & -8e^{-x/2} \\ + \; 24 & 16e^{-x/2} \\ - \; 0 & -32e^{-x/2} \end{array}$$

$$\overset{ch}{=} -2x^4e^{-x/2} - 16e^{-x/2}x^3 - 96x^2e^{-x/2} - 24\times 16\,x e^{-x/2} + 24\times(-32)e^{-x/2}$$

تمرین. فرض کنید $X_1, \dots, X_n \overset{iid}{\sim} Exp(\theta)$، ثابت کنید $Y = \sum_{i=1}^{n} X_i \sim \Gamma(n, \theta)$

نکته) حالت خاص از توزیع گاما که در آن $\alpha = n/2$ و $\theta = 1/2$ باشد به توزیع کی-دو با n درجه آزادی (پارامتر) معروف است و آن را با نماد $X \sim \chi^2_n$ نشان می‌دهیم.

★ ... اگر درجه آزادی ... فقط n تا بتواند تغییر کند و برای آن شرط نداشته باشیم...

یا

$X \sim \chi^2(n)$ کی-دو (chi-square)

۳- توزیع نمایی دو پارامتری.

در توزیع نمایی که تکیه گاه آن R^+ است. اگر بدانیم متغیر تصادفی مورد مطالعه حداقل می‌تواند مقدار μ را اختیار کند و سایر شرایط توزیع نمایی برقرار باشد، آنگاه متغیر مورد نظر دارای توزیع نمایی دو پارامتری با پارامترهای μ, θ است که دارای تابع چگالی احتمال زیر است:

$$f(x) = \begin{cases} \theta e^{-\theta(x-\mu)} & x > \mu \\ 0 & o.w. \end{cases}$$

۴- توزیع یکنواخت پیوسته.

گوئیم متغیر تصادفی X دارای توزیع یکنواخت روی بازه‌ی $(0, \theta)$ است هرگاه دارای تابع چگالی احتمال زیر

باشد:
$$f(x)=\begin{cases} 1/\theta & 0<x<\theta \\ 0 & o.w. \end{cases}$$

مصداق: اگر روی پاره خطی به طول θ، نقطه‌ای به تصادف انتخاب شود و متغیر تصادفی X طول نقطه‌ی انتخاب شده فرض کنید باشد، اگر احتمال انتخاب یا واقع شدن این نقطه در هر جایی هم طول از بازه‌ی $(0,\theta)$ همسان باشند، آنگاه X دارای توزیع یکنواخت روی بازه‌ی $(0,\theta)$ است که آن را با نماد $X \sim U(0,\theta)$ نشان می‌دهیم.

تمرین - تابع مولد گشتاور، امید ریاضی و واریانس توزیع $X \sim U(0,\theta)$ را حساب کنید. برگه

تمرین - روی پاره خطی به طول L نقطه‌ای به تصادف انتخاب می‌شود. چقدر احتمال دارد طول پاره خط بزرگتر حداقل دو برابر طول پاره خط کوچکتر باشد.

$$P(0<X\le L/3)+P(2L/3\le X<L)=\int_0^{L/3}\frac{1}{L}dx+\int_{2L/3}^{L}\frac{1}{L}dx = \dots = 2/3$$

مثال - از ایستگاه معینی از ساعت ۷ صبح به فاصله‌ی هر ۱۵ دقیقه اتوبوس عبور می‌کند. اگر یک عابر در زمانی بین ۷ تا ۷:۳۰ که به طور یکنواخت توزیع می‌شود به ایستگاه مراجعه کند، مطلوب است محاسبه‌ی احتمال اینکه

الف) کمتر از ۵ دقیقه برای اتوبوس منتظر بماند.

ب) حداقل ۱۲ دقیقه برای اتوبوس منتظر بماند.

X: زمان ورود فرد به ایستگاه: $X \sim U(0,30)$

(۷ ۷:۱۵ ۷:۳۰ ...)

الف) $P(10<X\le 15)+P(25<X\le 30)$

$$=\int_{10}^{15}1/30\,dx+\int_{25}^{30}1/30\,dx=\frac{5}{30}+\frac{5}{30}$$

$$=1/3$$

حل ب (me): $P(0<x\leq 3)+P(15<x\leq 18)=\int_0^3 \frac{1}{30}\,dx+\int_{15}^{18}\frac{1}{30}\,dx=\dots=\frac{1}{5}$

★ توزیع یکنواخت (گزوفند) $(\alpha=0,\ \theta=\beta$ جزوه$)$

$$f(x)=\begin{cases}\frac{1}{\beta-\alpha} & \alpha<x<\beta\\ 0 & o.w\end{cases}\Rightarrow M_x(t)=\frac{e^{t(\beta-\alpha)}-1}{t(\beta-\alpha)}$$

$$Ex=\frac{\alpha+\beta}{2}\qquad Var\,x=\frac{(\beta-\alpha)^2}{12}$$

هـ- توزیع نرمال.

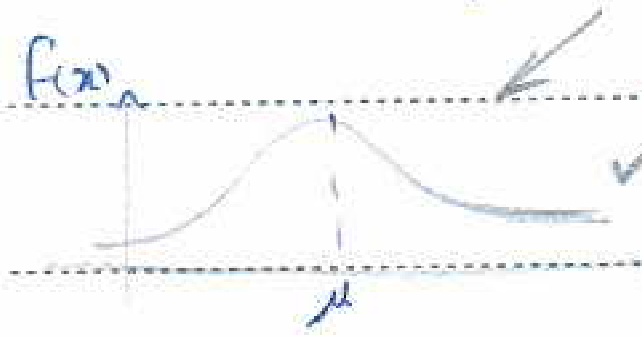
گوئیم متغیر تصادفی پیوسته ی X دارای توزیع نرمال با پارامترهای μ و σ^2 است، هرگاه دارای تابع چگالی احتمال زیر باشد.

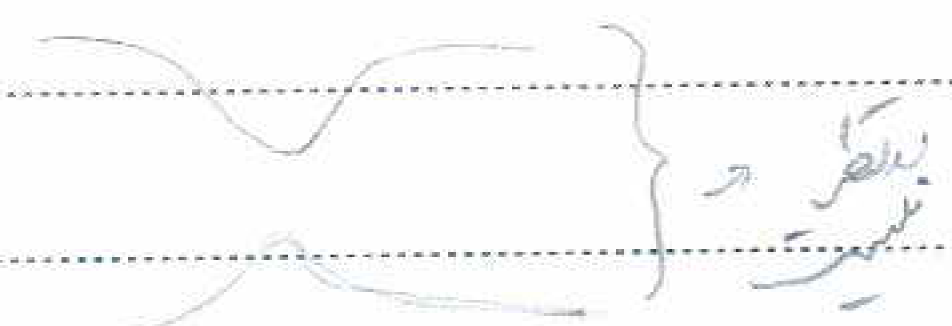
$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2\sigma^2}(x-\mu)^2}\qquad x\in\mathbb{R}$$

تمرین- تابع مولد گشتاور توزیع نرمال را محاسبه، امید ریاضی و واریانس آن را بدست آورید

$$M_x(t)=e^{\mu t+\frac{\sigma^2t^2}{2}}\qquad E(x)=\mu\qquad var(x)=\sigma^2$$

مصداق: اگر مجموعه ای از داده ها دارای رفتار احتمالی متقارن باشند (بعبارتی ضریب چاولگی آن ها صفر باشد) و منحنی مربوط به آن زنگوله ای شکل باشد، همچنین روابط زیر به طور همزمان برقرار باشند،





به مجموعه ی داده ها، داده های نرمال و به متغیر تصادفی بیین آن ها متغیر نرمال و به توزیع آن متغیر توزیع نرمال گفته می شود. و روابط ذکر شده عبارتند از:

1) $P(\mu-\sigma<X<\mu+\sigma)\approx 0.68$

2) $P(\mu-2\sigma<X<\mu+2\sigma)\approx 0.95$

3) $P(\mu-3\sigma<X<\mu+3\sigma)\approx 0.99$

ضریب چاولگی $sk=\frac{E(X-Ex)^3}{\sqrt{var(x)}^3}$

$P(-3<Z<3)\approx 0.99$ (نرمال استاندارد)

نکته) در توزیع نرمال اگر $\mu=0$ و $\sigma^2=1$ به آن، توزیع نرمال استاندارد می‌گوییم.

نکته ۲) اگر x دارای توزیع نرمال با پارامترهای μ و σ^2 باشد، آن را با نماد زیر نشان می‌دهیم: $X\sim N(\mu,\sigma^2)$

بنابراین $N(0,1)$ نشان دهنده‌ی توزیع نرمال استاندارد است.

هر ترکیب خطی x نیز دارای توزیع نرمال است؟

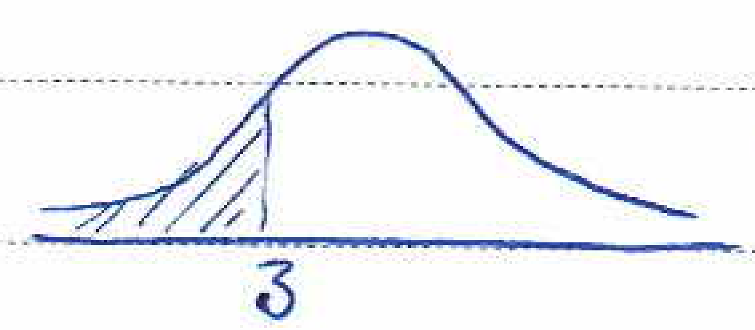
نکته ۳) اگر $X\sim N(\mu,\sigma^2)$، آنگاه $Z=\frac{X-\mu}{\sigma} \overset{\text{قابل اثبات}}{\sim} N(0,1)$.

★ نکته) اگر $Z\sim N(0,1)$ آنگاه: $P(Z\le z)=\int_{-\infty}^{z}\frac{1}{\sqrt{2\pi}}e^{-t^2/2}\,dt$

که انتگرال فوق با روش‌های عددی قابل محاسبه است. احتمال فوق به ازای مقادیر مختلف z محاسبه و در جدولی به نام "جدول توزیع نرمال" گردآوری شده است.

z	0.00	0.01	0.02	...	0.09
0.0					
0.1					
0.2			$\square$		
⋮					
0.9					
1.0					
1.1					
⋮					

رسم به راهنمای جدول:



یا $P(Z\le z)$

$P(Z\le 0.22)=?=\square$

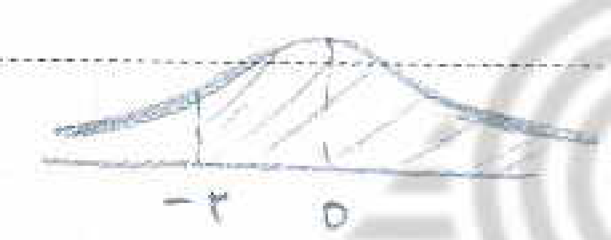
مثال. فرض کنید سود روزانه‌ی شرکت خاصی دارای توزیع نرمال با میانگین ۱۰۰ و واریانس ۲۵ باشد، احتمال اینکه سود امروز شرکت حداقل ۹۰ باشد چه قدر است؟

$X\sim N(100,25)$ سود شرکت

$P(X\ge 90) \overset{*}{=} P\left(\underbrace{\frac{X-100}{5}}_{Z}\ge\frac{90-100}{5}\right)=P(Z\ge -2)=\begin{cases}1-P(Z\le -2)\\ \text{یا}\\ P(Z\le 2)\end{cases}=$

*تبدیل به $N(0,1)$

به علت تقارن توزیع نرمال



Subject:
Year. Month. Date. ( )

★ برای طبیعتی که نمی تواند منفی باشد هم می توان توزیع نرمال را در نظر گرفت.

$= 0/9772$

تمرین – اگر Z دارای توزیع نرمال $Z \sim N(0,1)$ باشد آنگاه: $Z^2 \sim \chi^2(1)$.

$e^{\frac{t^2}{2}}$

$E(e^{t\sum_{i=1}^{n} x_i}) = \dots = \prod_{i=1}^{n} E(e^{tx_i})$

$= \left(\frac{1}{1-2t}\right)^{n/2}$

$(\checkmark = \left(\frac{1/2}{1/2-t}\right)^{n/2})$

تمرین ۲ – اگر $X_1, \dots, X_n \overset{iid}{\sim} \chi^2(1)$ آنگاه $\sum_{i=1}^{n} X_i \sim \chi^2(n)$.

نکته) اگر $Z_1, \dots, Z_n \overset{iid}{\sim} N(0,1)$ آنگاه $\sum_{i=1}^{n} Z_i^2 \sim \chi^2(n)$.

قضیه حد مرکزی. فرض کنید $X_1, \dots, X_n$ متغیرهای تصادفی مستقل و هم توزیع از توزیعی با میانگین μ و واریانس σ^2 باشند، در این صورت اگر حجم نمونه بسیار بزرگ باشد (به اندازه کافی بزرگ باشد) یعنی $n \to +\infty$

آنگاه: $\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \overset{\text{تقریبا}}{\approx} N(0,1)$ $(*)$ که در آن: $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} x_i$,

$\xrightarrow[n\to\infty]{}$

(★ $\bar{X}$ میانگین نمونه‌ی جامعه ولی μ میانگین جامعه.)
متغیر تصادفی پارامتر

رابطه‌ی $(*)$ را می توانیم به صورت زیر بازنویسی کنیم:

$$\frac{\sum_{i=1}^{n} x_i - n\mu}{\sigma\sqrt{n}} \xrightarrow[n\to\infty]{} N(0,1)$$

از طرفی می دانیم اگر $X_1, \dots, X_n$ متغیرهای تصادفی مستقل و هم توزیع با میانگین μ و واریانس σ^2 باشند، آن گاه:

$$E\left(\sum_{i=1}^{n} x_i\right) = \sum_{i=1}^{n} \overbrace{E x_i}^{\mu} = n\mu$$

$$Var\left(\sum_{i=1}^{n} x_i\right) = \sum_{i=1}^{n} Var(X_i) = n\sigma^2$$

بنابراین با فرض $Y = \sum_{i=1}^{n} X_i$ رابطه‌ی $(*)$ را می توان مجدداً به صورت زیر نوشت:

$$\frac{Y_n - E(Y_n)}{\sqrt{Var(Y_n)}} \xrightarrow[n\to\infty]{} N(0,1) \quad (**)$$

$Var(ax+by) = a^2 Var(x) + b^2 Var(y) + 2ab\,Cov(x,y)$ → اگر X و Y مستقل باشند $= 0$

فصل پنجم

نمونه‌گیری و برآورد: در یک جامعه‌ی مورد بررسی معمولاً برخی از شاخص‌های جامعه که به آن‌ها پارامتر می‌گوییم مجهول هستند. برای اطلاع از وضعیت این پارامترها بایستی نمونه‌ای تصادفی از آن جامعه انتخاب شود و بر اساس نمونه‌ی به دست آمده آن پارامترهای مجهول برآورد شود. برخی از شاخص‌های جامعه ← پارامتر

تعریف پارامتر: مشخصه‌ای از جامعه است که مقداری ثابت اما مجهول است. (مشخصه کمیتی است که خصوصیتی از جامعه را مشخص می‌کند مثل واریانس، میانگین و ...)

تعریف آماره: مشخصه‌ای از نمونه است که مقدار آن از نمونه‌ای به نمونه‌ای دیگر تغییر می‌کند و از آنجایی که نمونه‌ها به تصادف انتخاب می‌شوند تغییرات آماره نیز تصادفی است یعنی یک آماره یک متغیر تصادفی است که به پارامتر مجهول جامعه بستگی ندارد.

تعریف نمونه‌گیری: به عمل انتخاب بخشی از جامعه طبق اصول و ضوابطی خاص نمونه‌گیری گفته می‌شود. و به آن بخش از جامعه که انتخاب می‌شود یک نمونه می‌گوییم. نمونه‌گیری انواع متفاوتی دارد: ۱) نمونه‌گیری تصادفی ساده
۲) نمونه‌گیری طبقه‌ای ۳) نمونه‌گیری خوشه‌ای ۴) نمونه‌گیری سیستماتیک یا دوره‌ای

تعریف جامعه آماری: مجموعه‌ای از افراد یا اشیاء که در یک تحقیق مورد بررسی قرار می‌گیرند را یک جامعه آماری گوییم.

واحد آماری: به هر یک از اعضای یک جامعه آماری گفته می‌شود.

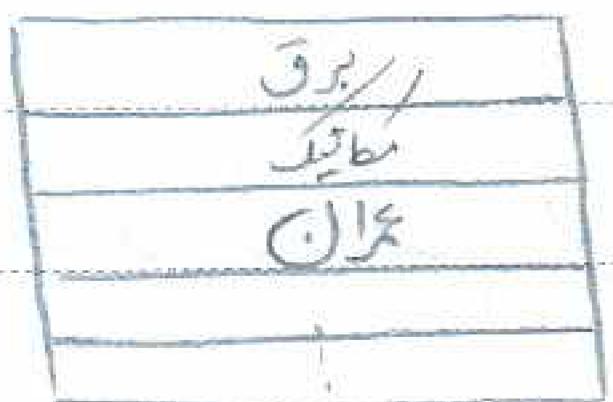
Subject:
Year. Month. Date. ( )

در یک تحقیق آماری علاقه مندیم یک یا چند صفت جامعه را مورد بررسی قرار دهیم. صفات مورد بررسی به دو دسته تقسیم می شوند: ۱) صفات مشخصه یا ثابت: صفاتی هستند که در بین کلیه ی واحدهای آماری یکسان می باشند. ۲) صفات آماری یا متغیر تصادفی: که از واحدی به واحد آماری دیگر تغییر می کنند.

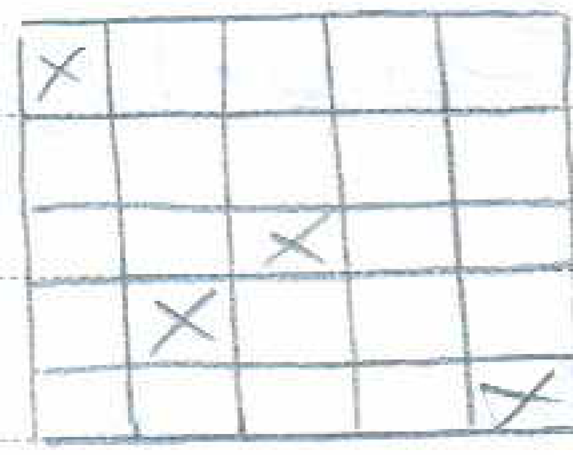
چارچوب آماری: به لیست تمام افراد (واحدهای) آماری یک جامعه ی آماری اطلاق می شود.

$N=1000$

۱) نمونه گیری تصادفی ساده: مثلاً در یک کلاس هر دانشجو یک شماره دارد، از ۱۰۰۰ عدد ۱۰۰ عدد تصادفی ایجاد می کنیم و صفت ها را بررسی می کنیم به دو روش * باجایگذاری و بدون جایگذاری.



۲) نمونه گیری طبقه ای: در حالت قبل ممکن است هر ۱۰۰ نفر رشته ی برق باشند پس دانشکده را طبقه بندی می کنیم و از این روش تعدادی از رشته برق و ... در نظر می گیریم.



۳) نمونه گیری خوشه ای: جامعه را دسته بندی می کنیم، چند تا دسته را تصادفی انتخاب می کنیم پس این دسته ها را به طور کامل سرشماری می کنیم داخل هر دسته واحدهای آماری وجود دارد

۴) نمونه گیری سیستماتیک یا دوره ای: از بین چند نفر که رد می شوند یک نفر را انتخاب فرم می دهند مثلاً هر ۳ نفر را یک نفر

$X_1 \rightarrow$ مقدار صفت (مورد مطالعه برای نفر اول) متناظر با:

* باجایگذاری: نمونه را گرفتیم واحد آماری را به جامعه برمی گردانیم نتیجه آزمایش اول "تاثیری در آزمایش دوم ندارد".

* بدون جایگذاری: نمونه ها از هم مستقل نمی شوند. اگر حجم N (N بیشتر از ۲۰ برابر n) خیلی بیشتر از n باشد می توان این فرض را کرد که اگر بدون جایگذاری انجام شود مستقل می شود. * اگر مستقل باشد کواریانس ها صفر می شود.

تذکر * اگر N (حجم جامعه) بیشتر از ۲۰ برابر n (حجم نمونه) باشد می توان فرض کرد که $X_1, \ldots, X_n$ تقریباً مستقل از هم اند.

قرارداد: وقتی گوییم $X_1, \dots, X_n$ یک نمونه تصادفی از جامعه‌ی مورد مطالعه است یعنی $X_1, \dots, X_n$ متغیرهای تصادفی iid از آن جامعه هستند. !!

تعریف. اگر کلیه‌ی n تایی‌های ممکن از یک جامعه به حجم N شانس مساوی برای رخداد داشته باشند به آن نمونه‌گیری، "نمونه‌گیری تصادفی ساده" می‌گوییم.

$\{a, b, c\}$

$\{a,b\}\{b,a\}$
$\{a,c\}\{c,a\}$
$\{b,c\}\{c,b\}$

* با جایگذاری و با رعایت ترتیب

برآوردگر: آماره‌ای است که برای تخمین مقدار واقعی پارامتر جامعه استفاده می‌شود و به مقدار آن برآورد پارامتر (مشخصه‌ای از نمونه) گفته می‌شود.

نکته) پارامترهای مهم یک جامعه عبارتند از: ۱- میانگین جامعه (μ) ۲- واریانس جامعه (σ^2)
۳- نسبت در جامعه (P)

اگر $X_1, \dots, X_N$ مقادیر صفت مورد بررسی در جامعه باشند آنگاه:

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i \qquad \sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2$$

$$P = \frac{A}{N}$$ ⟶ (A تعداد واحدهای آماری که دارای خصوصیت ویژه‌ای هستند.)

نکته) مهم‌ترین آماره‌های نمونه عبارتند از: ۱) میانگین نمونه $(\bar{X})$ ۲) واریانس نمونه (S^2) ۳) نسبت در نمونه $(\hat{P})$

Subject:
Year. Month. Date. ( )

اگر $X_1, \dots, X_n$ مقادیر صفت مورد بررسی در نمونه انتخاب شده باشند آنگاه:

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 \qquad \bar{p} = \frac{a}{n}$$

(a مقدار واحدهایی در نمونه که دارای خصوصیت ویژه ای هستند.)

ویژگی های یک برآوردگر خوب: ۱) نااریبی ۲) سازگاری ۳) کارایی

فرض کنید θ پارامتر مورد مطالعه و $\hat{\theta}$ (برآوردگر θ) برآوردگر آن باشد در این صورت معیارهای فوق را به شکل زیر تعریف می کنیم:

۱) نااریبی: $\hat{\theta}$ را یک برآوردگر نااریب برای θ گوییم هرگاه: $E(\hat{\theta}) = \theta$.

۲) سازگاری: اگر حجم نمونه خیلی زیاد باشد. ترالم / برآوردگر سازگار برآوردگری است که با افزایش حجم نمونه مقدار برآوردگر به مقدار واقعی پارامتر جامعه نزدیک تر می شود به عبارتی اگر $\hat{\theta}_n$ برآوردگر به دست آمده از نمونه ای به حجم n باشد آنگاه $\hat{\theta}_n$ را برای θ سازگار گوییم هرگاه:

$$\lim_{n\to\infty} P(|\hat{\theta}_n - \theta| > \varepsilon) = 0 \qquad \lim_{n\to\infty} P(|\hat{\theta}_n - \theta| < \varepsilon) = 1$$

✓ $\hat{\theta}_n$ و θ خیلی به هم دیگر نزدیک باشند وقتی که حجم نمونه زیاد باشد.

۳) کارایی: کارایی یک مفهوم نسبی است و از بین دو برآوردگر نااریب آنکه واریانس کمتری داشته باشد، کاراتر است.

معیار دیگری که می توان برای مقایسه ی بین برآوردگرها استفاده نمود، معیار MSE (میانگین مربع خطا)

Min square Error

می‌باشد که به صورت زیر تعریف می‌شود:

فرض کنید $\hat{\theta}_n$ یک برآوردگر برای θ باشد، در این صورت $MSE(\hat{\theta}_n)$ به صورت زیر تعریف می‌شود:

$$MSE(\hat{\theta}_n)=E(\hat{\theta}_n-\theta)^2$$

نکته) در بین برآوردگرهای مختلف آن که MSE کمتری دارد بهتر است، به عبارتی کاراتر است.

تعریف eff. فرض کنید T_1 و T_2 دو برآوردگر برای θ باشند در این صورت:

$$eff(T_1,T_2)=\frac{\frac{1}{MSE(T_1)}}{1/MSE(T_2)}=\frac{MSE(T_2)}{MSE(T_1)}$$

بنابراین:

$$eff(T_1,T_2)\begin{cases}>1 \Rightarrow & T_1 \text{ از } T_2 \text{ کاراتر است.}\\ =1 \Rightarrow & \text{کارایی } T_1 \text{ و } T_2 \text{ یکسان است.}\\ <1 \Rightarrow & T_2 \text{ کاراتر از } T_1 \text{ است.}\end{cases}$$

تعریف. میزان اریبی برآوردگر T برای پارامتر θ را به شکل زیر تعریف می‌کنیم: $b(T)=E(T)-\theta$

نکته) اگر T برای θ نااریب باشد آنگاه: $b(T)=0$.

تمرین. ثابت کنید $\star\ MSE(T)=Var(T)+b^2(T)$ (ک.س)

چپ: $=ET^2-(ET)^2+(ET)^2+\theta^2-2ET\theta$

راست: $=E(T^2)+\theta^2-2E_T\theta=E(T-\theta)^2$

نتیجه) اگر T برای θ نااریب باشد $MSE(T)=Var(T)$

مثال. فرض کنید $X_1, X_2, \dots, X_n$ یک نمونه‌ی تصادفی از توزیع $N(\mu,\sigma^2)$ باشد. برآوردگرهای

$$\equiv X_1, X_2, \dots, X_n \overset{iid}{\sim} N(\mu,\sigma^2)$$

★ کارایی MSE ٪ بهتر از نااریبی است.

E ... به دست آوردیم

✓ پایان ترم ۸-۷ سوال ← استدلال بحث (مصداق ۴★ توزیع توام)

$T_1 = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ ، $T_2 = \frac{X_1 + X_n}{2}$ را برای μ در نظر بگیرید. آن‌ها را از لحاظ نااریبی و کارایی مقایسه کنید. حل:

$$E(T_1) = E\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) = \frac{1}{n}\sum_{i=1}^{n} \underbrace{E(x_i)}_{\mu} = \mu \rightarrow$$ T_1 برای μ نااریب است

$$E(T_2) = E\left(\frac{X_1 + X_n}{2}\right) = \frac{1}{2}(\underbrace{Ex_1}_{\mu} + \underbrace{Ex_n}_{\mu}) = \mu \rightarrow$$ T_2 برای μ نااریب است.

$$MSE(T_1) = Var(T_1) = Var\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) = \frac{1}{n^2} Var\left(\sum_{i=1}^{n} x_i\right) \overset{\text{مستقل}}{=} \frac{1}{n^2}\sum_{i=1}^{n} \underbrace{Var(x_i)}_{\sigma^2} = \frac{\sigma^2}{n}$$

$$MSE(T_2) = Var(T_2) = Var\left(\frac{X_1 + X_n}{2}\right) = \frac{1}{4}(\underbrace{Var X_1}_{\sigma^2} + \underbrace{Var X_n}_{\sigma^2}) = \frac{\sigma^2}{2}$$

بنابراین اگر $n>2$ برای آنگاه $MSE(T_1) < MSE(T_2)$ و T_1 از T_2 کاراتر است.

تمرین- فرض کنید $X_1, X_2, \dots, X_n \overset{iid}{\sim} b(p)$ ، دو برآوردگر $T_1 = \bar{X}$ و $T_2 = \frac{\sum x_i + 1}{n+1}$ را در نظر بگیرید، آن‌ها را از نظر نااریبی و کارآمدی مورد بررسی قرار دهید. ★ ← exam یک سوال مشابه

$$E(T_1) = E(\bar{x}) = E\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(x_i) = \mu \rightarrow$$ T_1 برای μ نااریب است

$$E(T_2) = E\left(\frac{\sum x_i + 1}{n+1}\right) = \frac{1}{n+1}\sum_{i=1}^{n} E(x_i + 1) = \frac{1}{n+1}\left(\sum_{i=1}^{n} \underbrace{Ex_i}_{\mu} + n\right) = \frac{\mu n + n}{n+1} = \frac{n}{n+1}(\mu + 1)$$

$$MSE(T_1) = Var(T_1) = Var\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} \overbrace{Var(x_i)}^{\sigma^2} \overset{\text{مستقل}}{=} \frac{\sigma^2}{n}$$

$$MSE(T_2) = Var(T_2) = Var\left(\frac{\sum x_i + 1}{n+1}\right) = \frac{1}{(n+1)^2} Var\left(\sum_{i=1}^{n}(x_i + 1)\right) \overset{\text{مستقل}}{=} \frac{1}{(n+1)^2}\sum_{i=1}^{n} Var(x_i + 1) \rightarrow$$ $Var(a) = 0$

$$= \frac{n\sigma^2}{(n+1)^2}$$

$$\frac{n^2}{(n+1)^2} < 1 \Rightarrow \frac{n}{(n+1)^2} < \frac{1}{n} \Rightarrow$$ برای $n>0$ ، T_2 کاراتر از T_1 است!

$\Rightarrow$ T_2 البته چون کارایی مهم‌تر از نااریبی است.

قضیه. اگر $\lim_{n\to\infty} MSE(\hat{\theta}_n) = 0$ آنگاه $\hat{\theta}_n$ برای θ سازگار است.

اثبات: بنابه نامساوی چبیشف:

$$P(|\hat{\theta}_n-\theta|>e) \leq \frac{E(\hat{\theta}_n-\theta)^2}{e^2}$$

$$= \frac{MSE(\hat{\theta}_n)}{e^2}$$

$$\Rightarrow \lim_{n\to\infty} P(|\hat{\theta}_n-\theta|>e) \leq \lim \frac{MSE(\hat{\theta}_n)}{e^2} = 0$$

$$\Rightarrow \lim_{n\to\infty} P(|\hat{\theta}_n-\theta|>e) = 0$$

روش های برآوردیابی.

۱) روش گشتاوری: (MME). ایده‌ی اصلی این روش این است که گشتاورهای نمونه بایستی بتواند گشتاورهای جامعه را توصیف کند.

- گشتاور مرتبه r ام جامعه را به صورت زیر تعریف می‌کنیم: $\mu_r = E(X^r)$ $(r=1,2,\ldots)$

به طور متناظر گشتاور مرتبه r ام نمونه را به شکل زیر تعریف می‌کنیم: $m_r = \frac{1}{n}\sum_{i=1}^{n} X_i^r$

برای استفاده از روش گشتاوری با شروع از $r=1$ گشتاور مرتبه r ام نمونه و جامعه را مساوی هم قرار می‌دهیم تا جایی که بتوانیم پارامترها را برحسب نمونه‌ی تصادفی بنویسیم.

مثال. فرض کنید $X_1,\ldots,X_n \overset{iid}{\sim} P(\lambda)$ برآوردگر گشتاوری λ را به دست آورید.

حل.

$$\left.\begin{array}{l} \mu_1 = E(X) = \lambda \\ m_1 = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X} \end{array}\right\} \Rightarrow \mu_1 = m_1 \qquad \boxed{\hat{\lambda} = \bar{X}}$$

Subject:
Year. Month. Date. ( )

مثال ۲. فرض کنید $X_1, \dots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$ ، برآوردهای گشتاوری μ و σ^2 را به دست آورید. حل.

$$\left.\begin{array}{l} \mu_1 = Ex = \mu \\ m_1 = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x} \end{array}\right\} \Rightarrow \hat{\mu} = \bar{x} \quad (1)$$

$$\left.\begin{array}{l} \mu_2 = E(x^2) = Var(x) + (Ex)^2 = \sigma^2 + \mu^2 \\ m_2 = \frac{1}{2}\sum_{i=1}^{n} x_i^2 = \overline{x^2} \end{array}\right\} \Rightarrow \mu_2 = m_2 \Rightarrow \sigma^2 + \mu^2 = \overline{x^2}$$

$$\overset{(1)}{\Rightarrow} \sigma^2 = \overline{x^2} - \bar{x}^2$$

$$\overset{*}{\Rightarrow} \boxed{\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

* از گشتاور مرتبه ۱ در مرتبه ۲ استفاده کردیم.

تمرین - فرض کنید $X_1, \dots, X_n \overset{iid}{\sim} U(0, \theta)$ برآورد گشتاوری θ را به دست آورید.

(حل

$$\left.\begin{array}{l} \mu_1 = \frac{\theta + 0}{2} \\ m_1 = \bar{x} \end{array}\right\} \Rightarrow \bar{x} = \frac{\theta}{2} \Rightarrow 2\bar{x} = \theta$$

۲) روش درست نمایی ماکزیمم (MLE). فرض کنید $X_1, \dots, X_n$ نمونه‌ای تصادفی از توزیعی با تابع چگالی (جرم) احتمال $f_\theta(x)$. در این صورت تابع درست نمایی پارامتر θ را بر حسب نمونه‌ی فوق

likely hood!

به شکل زیر می‌سازیم:

$$L(\theta) = \prod_{i=1}^{n} f_\theta(x_i) = f_\theta(\underbrace{x_1, \dots, x_n}_{\text{ثابت‌ها}})$$

متغیر تصادفی

برای به دست آوردن برآوردگر درست نمایی ماکزیمم یا MLE ی θ باید تابع $L(\theta)$ را ماکزیمم کنیم.

$f' = 0 \Rightarrow \dots$ ویژگی max یا min بودن

نکته) اگر x^* نقطه‌ی ماکزیمم کننده‌ی تابع f و g یک تابع صعودی باشد، آنگاه x^* نقطه‌ی ماکزیمم کننده‌ی تابع $g(f(x))$ نیز هست. و اگر g نزولی بود x^*، min کننده‌ی $g(f(x))$ می‌بود.

قبلی‌ها $f_\theta(\underbrace{x_1, \dots, x_n}_{\text{متغیرها}})$

الان‌ترها متغیر تصادفی ثابت

Subject: ۳۶
Year. Month. Date. ( )

برای همه توزیع‌ها چی؟!

نتیجه) نقطه ماکزیمم کننده‌ی تابع $L(\theta)$ با نقطه ماکزیمم کننده تابع $L(\theta)$ ln یکسان است.

مثال. فرض کنید $X_1, \dots, X_n \sim Exp(\theta)$ برآوردگر درست نمایی ماکزیمم θ را بدست آورید.

حل. $X \sim Exp(\theta) \Rightarrow f_\theta(x) = \begin{cases} \theta e^{-\theta x} & x>0 \\ 0 & o.w. \end{cases}$ — حتما توزیع گفته شده هم داده شود

$$\Rightarrow L(\theta) = \prod_{i=1}^{n} f_\theta(x_i) = \prod_{i=1}^{n} \theta e^{-\theta x_i} = \theta^n e^{-\theta \sum_{i=1}^{n} x_i}$$

$$\Rightarrow \ln(L(\theta)) = n\ln\theta - \theta\sum_{i=1}^{n} x_i \Rightarrow \frac{\partial \ln L(\theta)}{\partial\theta} = \frac{n}{\theta} - \sum_{i=1}^{n} x_i = 0 \Rightarrow \theta = \frac{n}{\sum_{i=1}^{n} x_i}$$

$$\Rightarrow \frac{\partial^2}{\partial\theta^2}\ln L(\theta) = \frac{-n}{\theta^2} < 0 \Rightarrow MLE(\theta) = \frac{n}{\sum_{i=1}^{n} x_i} = \frac{1}{\bar{X}}$$

$$\left.\begin{matrix} M_1 = 1/\theta \\ m_1 = \bar{X} \end{matrix}\right\} \Rightarrow \hat{\theta} = 1/\bar{X}$$

تمرین - برای همه‌ی تمرین‌ها برآوردگرهای MME و MLE را بدست آورده و مقایسه کنید. کارایی...

تمرین - فرض کنید $X_1, \dots, X_n \overset{iid}{\sim} U(0,\theta)$ برآوردگرهای گشتاوری و درست نمایی ماکزیمم θ را حساب کنید و آن‌ها را از لحاظ نااریبی، سازگاری، کارایی مقایسه کنید. مهم

$f_\theta(x) = \frac{1}{\theta} \quad 0 < x < \theta$ — در تمام مثال‌هایی که دامنه تغییرات به θ (پارامتر) مربوط است...

$$L(\theta) = \prod_{i=1}^{n} f_\theta(x_i) = \prod_{i=1}^{n} \frac{1}{\theta} = \frac{1}{\theta^n} \quad (0 < x_i < \theta)$$

مهم ملاحظه می‌شود که تابع $L(\theta)$ نسبت به θ یک تابع نزولی است بنابراین Max خود را در ابتدای بازه‌ی تغییرات θ اختیار می‌کند.

$$\left.\begin{matrix} X_1 < \theta \\ \vdots \\ X_n < \theta \end{matrix}\right\} \Rightarrow \max(X_1, \dots, X_n) < \theta$$

ادامه...

$\Rightarrow MLE\ \theta = \max(X_1,\dots,X_n) = y = T_1 \qquad MME = 2\bar{X} = T_2$

برای پایان ترم (use)

$E(y) = E(E(y|x))$

$f(x,y) = f(y|x)f(x) = f(x|y)f(y)$ — * + چندین مفهوم مصداق ...

$f(x,y) \qquad z = x/y$ — ✓ طریقه خواندن جدول توزیع نرمال

$E(\min(X_1 \dots X_n))$ — (توزیع نرمال آمده)

$\max(X_1 \dots X_n)$

گفته شده

واریانس دو متغیره - نامساوی چبیشف - مارکوف

قضیه حد مرکزی - توزیع ...

سیستم موازی — سیستم سری

اولین از کار افتادن طول عمر

$\max$(عمر مولفه) — آخرین از کار افتادن

$E(\min(X_1,\dots,X_n))$ سری

$\Rightarrow E(\text{سیستم موازی}) = E(\max\dots)$

نکته) اگر $X \sim B(n,p)$ که در آن n مقدار بسیار بزرگ $(n \geq 50)$ و p مقدار بسیار کوچک باشد، طوری که $\lambda = np$ آنگاه می‌توان توزیع دوجمله‌ای را با توزیع $P(\lambda)$ تقریب زد.

مثال. فرض کنید $X \sim B(200, 0.01)$ مطلوبست محاسبه‌ی دقیق و تقریبی $P(X=10)$؟

دقیق: $P(X=10) = \binom{200}{10}(0.01)^{10}(0.99)^{190}$

تقریبی: $X \approx P(2) \Rightarrow P(X=10) = \frac{e^{-2}\, 2^{10}}{10!}$

$\left(f(x) = \frac{\beta^{\alpha} x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)}\right) \quad x>0$

حل تمرین: $Z \sim N(0,1) \Rightarrow Z^2 \sim \chi^2_1$

$\chi^2_1 = \Gamma(1/2, 1/2)$

$X \sim \Gamma(\alpha,\beta) \Rightarrow M_X(t) = E(e^{tX}) = \int_0^\infty \frac{e^{tx}\beta^\alpha x^{\alpha-1}e^{-\beta x}}{\Gamma(\alpha)}dx = \int_0^\infty \frac{\beta^\alpha x^{\alpha-1}e^{-(\beta-t)x}}{\Gamma(\alpha)}dx =$

$= \frac{\beta^\alpha}{(\beta-t)^\alpha}\underbrace{\int_0^\infty \frac{(\beta-t)^\alpha x^{\alpha-1} e^{-(\beta-t)x}}{\Gamma(\alpha)}dx}_{=1} = \dots$

اثبات کنید: $\lim_{\substack{n\to\infty \\ p\to 0 \\ np\to\lambda}} \binom{n}{x}p^x(1-p)^{n-x} = \frac{e^{-\lambda}\lambda^x}{x!}\ ;\ \lambda = np$

$= \left(\frac{\beta}{\beta-t}\right)^{\alpha} \Rightarrow M_{\chi^2_1}(t) = \left(\frac{1/2}{\frac{1}{2}-t}\right)^{1/2} = \left(\frac{1}{1-2t}\right)^{1/2}$ ①

$M_{Z^2}(t) = E(e^{tZ^2}) = \int_{-\infty}^{\infty} e^{tz^2} f_Z(z)\, dz$

$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \; ; \; z \in \mathbb{R} \Rightarrow M_{Z^2}(t) = \int_{-\infty}^{\infty} e^{tz^2} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz$

$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(1-2t)}{2} z^2}\, dz$ ★

★ $= \sqrt{\frac{1}{1-2t}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{\frac{1}{1-2t}}\sqrt{2\pi}} e^{-\frac{z^2}{2/(1-2t)}}\, dz = \sqrt{\frac{1}{1-2t}}$ ② $\left(\sigma^2 = \frac{1}{1-2t}\right)$

می‌دانیم: $X \sim N(\mu, \sigma^2)$ $\quad f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

از ①، ② ملاحظه می‌شود که: $M_{Z^2}(t) = M_{\chi^2_1}(t)$ بنابراین بنابه یکتایی تابع مولد گشتاور نتیجه می‌شود که $Z^2 \sim \chi^2_1$

اثبات:

★ یادآوری: $X \sim N(\mu, \sigma^2)$, $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

$X \sim N(\mu, \sigma^2)$

$Z = \frac{X-\mu}{\sigma} \sim N(0,1) \Rightarrow X = \mu + \sigma Z$

$M_X(t) = E(e^{tX}) = E(e^{t(\mu+\sigma Z)}) = E(e^{\mu t} e^{\sigma t Z}) = e^{\mu t} \underbrace{E(e^{\sigma t Z})}_{M_Z(\sigma t)}$

$M_Z(t) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{tz} e^{-z^2/2}\, dz = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z^2 - 2tz)}\, dz$

$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(z-t)^2}{2} + \frac{t^2}{2}}\, dz$

$= e^{t^2/2} \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(z-t)^2}{2}}\, dz}$ → توزیع نرمال با $\sigma = 1$ و $\mu = t$ است... $= e^{t^2/2} \xRightarrow{t \to \sigma t} M_X(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}}$

$k(t) = \ln M_X(t) = \mu t + \frac{\sigma^2 t^2}{2}$

$k''(0) = \sigma^2$?

Subject : .... Date : / /

نکته ۲) حد مرکزی

$X_1,\dots,X_n \quad EX_i=\mu \ , \ var X_i=\sigma^2$

① $\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \longrightarrow N(0,1)$

$\dfrac{\frac{1}{n}\sum X_i-\mu}{\sigma/\sqrt{n}} \overset{②}{=} \dfrac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\,\sigma} \xrightarrow[n\to\infty]{} N(0,1)$

$Y=\sum_{i=1}^{n} X_i \overset{③}{\Rightarrow} \dfrac{Y-E(Y)}{\sqrt{var Y}} \longrightarrow N(0,1)$

مثال. فرض کنید $X\sim B(n,p)$ مطلوبست محاسبه تقریبی $P(Y=1)$ ؟ و $P(Y\le 1)$ ؟

می‌دانیم $Y=\sum_{i=1}^{n} X_i$ که در آن $X_1,\dots,X_n \overset{iid}{\sim} b(p)$

بنابراین براساس قضیه حد مرکزی:

$Z=\dfrac{Y-EY}{\sqrt{var Y}}=\dfrac{Y-np}{\sqrt{npq}} \xrightarrow{n\to\infty} \underbrace{N(0,1)}_{\text{پیوسته}}$

$P(Y=1)=P\left(Z=\dfrac{1-np}{\sqrt{npq}}\right)=0$ چون احتمال متغیر تصادفی پیوسته در یک نقطه صفر است.

$P(Y\le 1)=P\left(Z\le \underbrace{\dfrac{1-np}{\sqrt{npq}}}_{-\infty}\right)\overset{n\to\infty}{=}0$

$P(Y\ge 1)=P\left(Z\ge \dfrac{1-np}{\sqrt{npq}}\right)=1$

خیلی بزرگ (n=۵۰ در آمار بی‌نهایته!)

مثال ۴. فرض کنید $X\sim B(100,\ 0.5)$ مطلوبست محاسبه ی $P(X\le 20)$ ؟

$P(X\le 20)=P\left(\dfrac{X-EX}{\sqrt{varX}}\le \dfrac{20-EX}{\sqrt{varX}}\right)=P\left(Z\le \dfrac{20-50}{\sqrt{25}}\right)=P\left(Z\le \dfrac{-30}{5}\right)$

$=P(Z\le -6)\approx 0 \qquad Z\sim N(0,1)$

$P(X\le 60)=\cdots=P\left(Z\le \dfrac{60-50}{5}\right)=P(Z\le 2)=$ جدول طبق $\approx 0/98$

$E(T_2) = E(2\bar{X}) = E(\frac{2}{n}\sum_{i=1}^{n} x_i) = \frac{2}{n}\sum_{i=1}^{n} \underbrace{E(x_i)}_{\theta/2} = \theta$ (ادامه ی حل تمرین قبل نکته)

$\Rightarrow$ T_2 برای θ ناریب است.

$MSE(T_2) = var(T_2) = var(2\bar{X}) = Var(\frac{2}{n}\sum_{i=1}^{n} var(x_i)) \overset{\text{مستقل}}{=} \frac{4}{n^2}\sum_{i=1}^{n} \underbrace{var x_i} = \frac{\theta^2}{3n}$

$var X = \frac{\theta^2}{3} - (\frac{\theta}{2})^2 = \theta^2/12$

(حل کامل برای امتحان نوشته شود)

$E(T_1) = \int t f_{T_1}(t)\,dt$ $\quad f_X(x) = \begin{cases} 1/\theta & 0<x<\theta \\ 0 & o.w. \end{cases} \rightarrow f_{T_1}(t) = n f_X(t)[F_X(t)]^{n-1}$

$\underbrace{F_X(t)}_{P(X\le t)} = \int_0^t f_X(x)\,dx = \int_0^t 1/\theta\, dx = t/\theta \Rightarrow f_{T_1}(t) = \frac{n}{\theta}(\frac{t}{\theta})^{n} = \frac{nt^{n-1}}{\theta^n}$; $\underbrace{0<t<\theta}$

چون $0<x_i<\theta$ پس $0<\max x_i<\theta$

$\Rightarrow E(T_1) = \int_0^\theta t\frac{nt^{n-1}}{\theta^n}dt = \frac{n}{\theta^n}\int_0^\theta t^n dt = \frac{n\theta}{n+1}$

$\xrightarrow{\text{میزان اریبی } T_1} b(T_1) = E(T_1) - \theta = \frac{n\theta}{n+1} - \theta = \frac{-\theta}{n+1}$

$E(T_1^2) = \int_0^\theta t^2 f_{T_1}(t)\,dt = \int_0^\theta t^2\frac{nt}{\theta^n}dt = \frac{n}{\theta^n}\int_0^\theta t^{n+1}dt = \frac{n}{\theta^n}\,\frac{\theta^{n+2}}{n+2}$

$= \frac{n\theta^2}{n+2}$

$var(T_1) = \frac{n\theta^2}{n+2} - (\frac{n\theta}{n+1})^2 = \frac{n(n+1)^2 - n^2(n+2)}{(n+2)(n+1)^2}\theta^2$

$\Rightarrow MSE(T_1) = \left[\frac{n(n+1)^2 - n^2(n+2)}{(n+2)(n+1)^2} + \frac{1}{(n+1)^2}\right]\theta^2 = \left[\frac{n(n+1)^2 - (n+2)(n^2-1)}{(n+2)(n+1)^2}\right]\theta^2$

$= \dots$ (ساده کنید)

$\overset{?}{\Rightarrow} MSE(T_1) < MSE(T_2)$

در MSE میزان اریبی هم تاثیر دارد

ج همیشه نمی‌توان از $P(\frac{X-5}{3} > \frac{2-5}{3})$ استفاده کرد باید اول توجیه شود.

مثال. فرض کنید X دارای توزیع $X \sim \chi^2(n)$ باشد، مطلوبست محاسبه‌ی $P(X>2)$ با استفاده از تقریب نرمال:

حل. با توجه به اینکه $X = \sum_{i=1}^{n} X_i$ که در آن $X_1, \dots, X_n \overset{iid}{\sim} \chi^2(1)$ بنابراین با استفاده از قضیه حد مرکزی (یا با استفاده از رابطه (**)) می‌توانیم از تقریب زیر استفاده کنیم:

$$\frac{X-EX}{\sqrt{Var(X)}} = \frac{X-n}{\sqrt{2n}} \xrightarrow[n\to\infty]{} N(0,1)$$

در نتیجه:

$$P(X>2) = P\left(\frac{X-n}{\sqrt{2n}} > \frac{2-n}{\sqrt{2n}}\right) = P\left(Z > \underbrace{\frac{2-n}{\sqrt{2n}}}_{n\to\infty}\right) = P(Z > -\infty) = 1$$

+ با پاورقی

محاسبه دقیق مثال اول صفحه:

$$\int_2^\infty \frac{x^{n/2-1}\, e^{-x/2}}{2^{n/2}\,\Gamma(n/2)} = \dots \longrightarrow \dots$$

برنولی : X : تعداد موفقیت در ۱ آز $\quad p^x(1-p)^{1-x}$

دوجمله‌ای : X : تعداد موفقیت در n آز $\quad \binom{n}{x} p^x (1-p)^{n-x}$

هندسی : X : تعداد آز لازم تا حصول اولین موفقیت $\quad p(1-p)^{x-1}$

پاسکال : X : تعداد آز لازم تا حصول r امین موفقیت $\quad \binom{x-1}{r-1} p^r q^{x-r}$

پواسون : $\dfrac{e^{-\lambda}\lambda^x}{x!}$

(فرمول) r امین گشتاور حول مبدا توزیع گاما : $\mu'_r = \dfrac{(1/\theta)^r\,\Gamma(\alpha+r)}{\Gamma(\alpha)}$