آمار

Subject: ۹۶/۱۱/۱۷

جلسه اول

احتمال

آزمایش تصادفی: آزمایشی است که در مورد خروجی قطعیت نداریم.

مثال: پرتاب یک سکه

شیر
خط

فضای نمونه (Ω): مجموعه‌ی خروجی‌های ممکن یک آزمایش تصادفی.

event یا رخداد یا پیشامد: یک زیرمجموعه از فضای نمونه‌ای است

مثال: انداختن یک تاس

1	2
3	4
5	6

event ←

* اگر یکی بود رخداد ساده است

احتمال: یک مقدار عددی که به یک رخداد نسبت داده می‌شود.

اصول احتمال کولموگروف:

Axioms

اگر یک فضای نمونه‌ای با نام Ω داشته باشیم و رخدادهای A و B در این فضا

منظم کنیم:

(1) $P(A) \geq 0 \quad \forall A \subset \Omega$

IDEA

Subject:

(۲) $P(\Omega)=1$ normalization (Ω شامل همه‌ی

خروجی‌های ممکن آزمایش)

(۳) برای هر دو رخداد (زیرمجموعه) A , B که A , B ناسازگار باشند

$$P(A\cup B)=P(A)+P(B) \qquad (A\cap B=\varnothing)$$

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

$$P(A')=P(A^c)=1-P(A)$$

$$P(A\cup A')=P(\Omega)=1=P(A)+P(A')$$

گسترش اصل (۳) : اگر تعداد محدود (متناهی) از رخدادهای دو به دو ناسازگار داشته باشیم

خواهیم داشت : $P(A_1\cup A_2 \dots \cup A_n)=P(A_1)+\dots+P(A_n)$

قانون احتمال توزیع گسسته یکنواخت :

اگر احتمال همه خروجی‌های یک آزمایش تصادفی یکسان باشد ، احتمال رخداد یک زیرمجموعه

A از فضای نمونه برابر است با :

$$P(A)=\frac{n_A}{n_\Omega}$$

Subject:

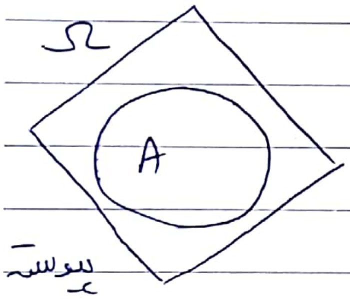
مثال: اگر ۲ تاس پرتاب کنیم، احتمال اینکه مجموع دو تاس برابر ۷ باشد چقدر است؟

$|\Omega| = ۶\times۶ = ۳۶$

(۱,۶) (۶,۱) (۴,۳) (۳,۴) (۵,۲) (۲,۵)

$\frac{nA}{n\Omega} = \frac{۶}{۳۶}$

قانون احتمال توزیع پیوستهٔ یکنواخت:

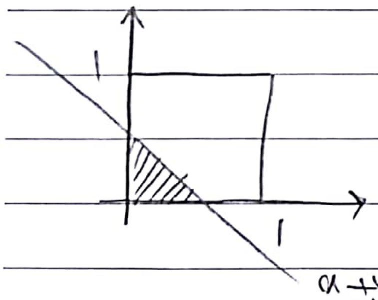


همهٔ زیربازه‌های مساوی با همدیگر احتمال یکسان را داشته باشند.

$\Downarrow$

$P(A) = \frac{\text{مساحت}\ (A)}{\text{مساحت}\ (\Omega)}$

مثال: اگر به یک مربع یک در یک، سوزنی پرتاب کنیم چقدر احتمال دارد که مجموع طول و عرض نقطهٔ اصابت کمتر از $\frac{1}{2}$ باشد؟



$P(A) = \frac{1}{8}$

اصل تقسیم‌پذیری کولموگروف (برای حالت ناشمارا):

برای هر دنباله‌ای از رخدادها دوبه‌دو ناسازگار، تعدادشان شمارا باشد و قابل ترتیب‌بندی باشد.

Subject:

(حالت‌های صعودی یا نزولی) خواهیم داشت:

$$P(B_1 \cup B_2 \dots \cup B_n) = \sum_{i=1}^{\infty} P(B_i) = \lim_{n \to \infty} \sum_{i=1}^{\infty} P(B_i)$$

مثال: یک سکه را آنقدر پرتاب می‌کنیم تا اولین بار H (head) بیاید.

احتمال اینکه سکه را به دفعات زوج پرتاب کنیم، چقدر است؟

غلط

$\cancel{\frac{1}{2}}$ → حالات زوج و فرد دلیل

دفعه ۲

$$\underbrace{\frac{1}{2}}_{T} \times \underbrace{\frac{1}{2}}_{H} + \frac{1}{2} \times \underbrace{\frac{1}{2} \times \frac{1}{2}}_{T} \times \frac{1}{2} + \dots$$

$$P(\text{دفعات زوج}) = \frac{\frac{1}{4}}{1-\frac{1}{4}} = \frac{1}{3}$$

$$P(\text{دفعات فرد}) = \frac{\frac{1}{2}}{1-\frac{1}{4}} = \frac{2}{3}$$

$\{H\}$ H می‌تونه دفعه‌ی اول بیاد

$$\boxed{\begin{array}{l} \{H\} \quad \{T,H\} \qquad \text{دفعه‌ی دوم} \\ \{T,T,H\} , \{T,T,T,H\} \end{array}}$$

احتمال tail آمدن $\underbrace{\underset{T}{\frac{1}{2}} \times \underset{H}{\frac{1}{2}}}_{\text{دفعه‌ی دوم}} + \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} + \dots$

$$= \frac{1}{2^2} + \frac{1}{2^4}$$

IDEA

قدر نسبت $= \frac{1}{4}$

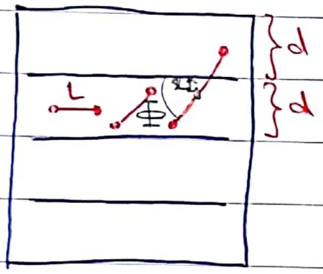
Subject:

رخدادهای ناسازگار یعنی اگر i دفعه بدوم، احتمالا دیگر دفعه j ام اتفاق نمی‌افتد.

رخدادهای خاص، رخدادها هستند.

جلسه دوم   ۹۴/۱۱/۲۴

مثال: یک صفحه با خطوط موازی به فاصله d در نظر بگیرید. یک سوزن به طول L روی این صفحه پرتاب می‌شود اگر $L<d$، با چه احتمالی سوزن یک خط را قطع می‌کند؟



x, Φ

x: فاصله مرکز سوزن از نزدیک‌ترین خط

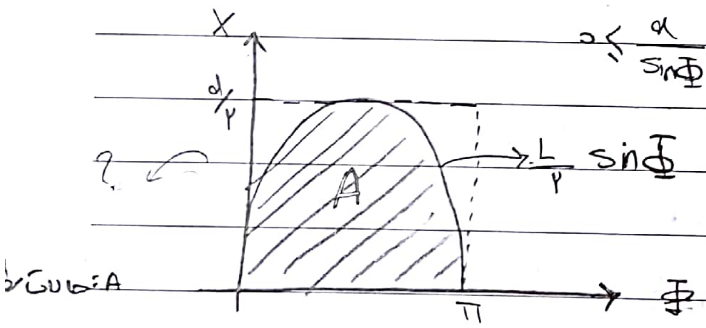
$$\boxed{\begin{array}{l} 0 < \Phi < \pi \\ 0 < x < \frac{d}{2} \end{array}}$$

فضای نمونه

حالات مطلوب

$$0 < \frac{x}{\sin\Phi} < \frac{L}{2}$$

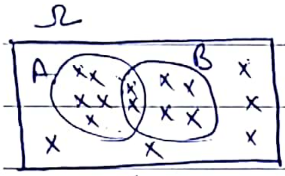
$$\boxed{\begin{array}{l} 0 < x < \frac{L}{2}\sin\Phi \\ 0 < \Phi < \pi \end{array}}$$



$$P(\text{برخورد سوزن}) = \frac{\text{مساحت } A}{\text{مساحت کل}} = \frac{\int_0^{\pi} \frac{L}{2}\sin\Phi \, d\Phi}{\pi \times \frac{d}{2}}$$

$$= \frac{\frac{L}{2}\left(-\cos\Phi \Big|_0^{\pi}\right)}{\frac{d\pi}{2}} = \frac{2L}{d\pi}$$

Subject:



احتمال شرطی:

$P(A)=\frac{7}{14}$

$P(B)=\frac{4}{14}$

$$P(A\mid B)=\frac{P(A\cap B)}{P(B)}=\frac{1}{3}$$

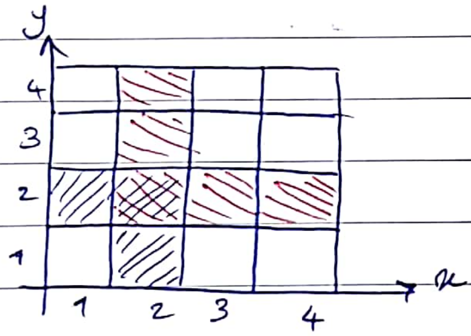
هر سه حالت ممکن است.
$$\begin{cases} P(A)=P(A\mid B)\\ P(A)>P(A\mid B)\\ P(A)<P(A\mid B)\end{cases}$$

مثال: دو تاس چهاروجهی را پرتاب می‌کنیم:

الف) احتمال اینکه max اعداد آمده برابر ۲ باشد؟ $\frac{3}{14}$

$P(\max(x,y)=2)=?$

ب) احتمال اینکه max اعداد آمده برابر ۲ باشد به شرط اینکه min هم برابر ۲ باشد؟



$$P(\underbrace{\max(x,y)=2}_{A}\mid \underbrace{\min(x,y)=2}_{B})=$$

$$\frac{P(A\cap B)}{P(B)}=\frac{1/14}{5/14}=\frac{1}{5}$$

مثال: برای یک رادار اطلاعات زیر را در اختیار داریم:

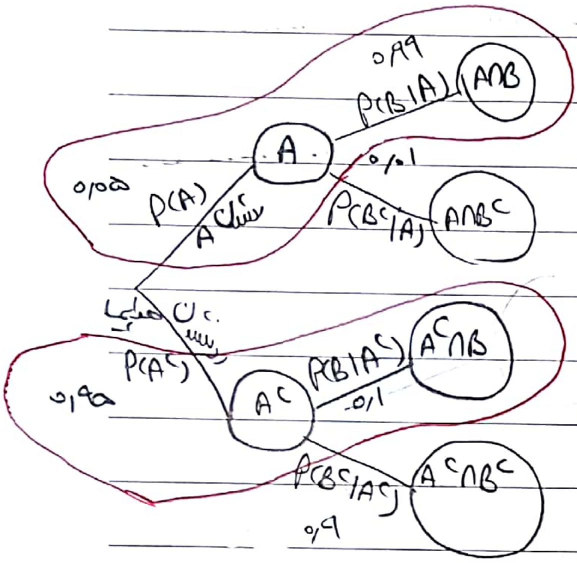
- همواره با احتمال ۰٫۰۵ امکان عبور هواپیما از بالای رادار وجود دارد. $P(A)=0.05$
- در صورت رد شدن هواپیما این رادار با احتمال ۰٫۹۹ هشدار می‌دهد. $P(B\mid A)=0.99$
- در صورت رد نشدن هواپیما این رادار با احتمال ۰٫۱ هشدار می‌دهد. $P(B\mid A^c)=0.1$

آیا این رادار عملکرد قابل قبولی دارد؟

Subject:

B: هشدار دادن رادار　　　　A: رخداد ردشدن هواپیما

$P(A|B)$ = چقدر احتمال دارد هواپیما ردشود به شرط اینکه رادار هشدار دهد



$$P(A|B) = \frac{P(A\cap B)}{P(B)} = ۰٫۳۴$$

$$P(B) = \underbrace{P(A\cap B)}_{P(A)\,.\,P(B|A)} + \underbrace{P(A^c\cap B)}_{P(A^c)\,.\,P(B|A^c)}$$

$$P(B) = ۰٫۰۵ \times ۰٫۹۹ + ۰٫۹۵ \times ۰٫۱$$

دو رخداد A, B از هم مستقل هستند اگر و فقط اگر $P(A\cap B)=P(A).P(B)$.

$$P(A) = P(A|B) \qquad P(B)\neq 0$$

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

A, B ناسازگار باشند آیا A, B از هم مستقل هستند؟

$P(A\cap B) = 0$　　ناسازگاری

$P(A\cap B) = P(A).P(B) = 0$

دو رخداد A, B ناسازگار مستقل هستند اگر حداقل احتمال رخداد یکی از آنها صفر باشد.

Subject:

دو تاس پرتاب می کنیم احتمال اینکه مجموع ۷ بیاید ؟ $P(A) = \frac{6}{36} = \frac{1}{6}$

$P(B)$

الف) آیا رخداد مجموع ۷ و رخداد اینکه تاس اول برابر ۳ باشد از هم مستقل هستند ؟

باید ثابت کنیم $P(A \cap B) \overset{?}{=} P(B) . P(A)$ $P(B) = \frac{1}{6}$

$\frac{1}{36} \overset{\checkmark}{=} \frac{1}{6} \times \frac{1}{6}$

بله مستقل است $P(A \mid B) = \frac{1}{6}$

جلسه سوم ۹۹,۱۱,۲۹

دو تاس پرتاب می کنیم

واقعه A : احتمال اینکه مجموع ۷ بیاید. $P(A) = \frac{1}{6}$

واقعه B : تاس اول ۳ بیاید $P(B) = \frac{1}{6}$

واقعه C : فقط یکی از تاس ها ۳ بیاید. $P(C) = \frac{10}{36}$

A , B از هم مستقل هستند $P(A \cap B) = P(B) . P(A)$

$\frac{1}{36} = \frac{1}{6} \times \frac{1}{6}$

A , C از هم مستقل نیستند $P(A \cap C) = P(A) . P(C)$

$\frac{2}{36} \neq \frac{1}{6} \times \frac{1}{36}$

IDEA

Scanned by CamScanner

Subject:

مثال: یک کشوری احتمال ابتلا به سرطان خون ۱٪ است. از بین افرادی که سرطان خون دارند، نتیجه آزمایش خون ۸۰٪ افراد مثبت است. از بین افرادی که سرطان ندارند، آزمایش خون آنها ۹۰٪ منفی است. آیا پزشک با دیدن آزمایش خون + می‌تواند به داشتن سرطان اطمینان دهد؟

واقعه A: داشتن سرطان

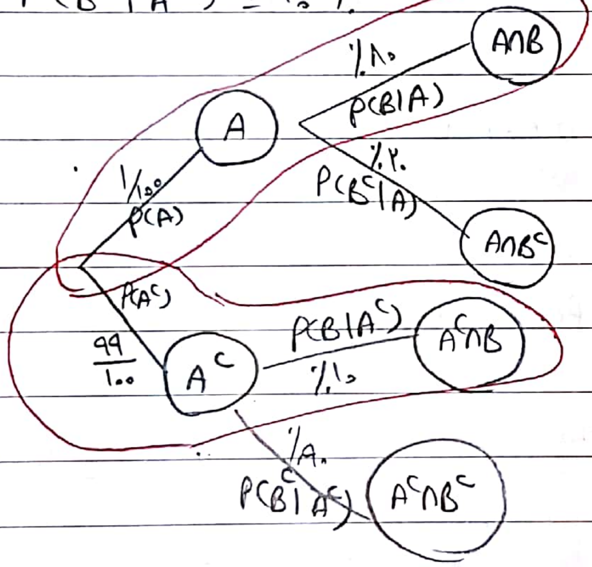
$P(A) = \frac{1}{100}$

واقعه B: آزمایش خون +

$P(B \mid A) = 80\%$

$P(B^c \mid A^c) = 90\%$

$P(B \mid A^c) = 10\%$



قانون بیز

$$P(A \mid B) = \frac{P(A\cap B)}{P(B) = P(B\cap A) + P(B\cap A^c)}$$

$$= \frac{P(A\cap B)}{P(A).P(B\mid A)}$$

$$= \frac{P(A\cap B)}{P(A).P(B\mid A) + P(A^c).P(B\mid A^c)}$$

$$= \frac{0.01 \times 0.8}{0.01 \times 0.8 + 0.99 \times 0.1} = \frac{8}{107}$$

Subject:

$$P(B) = P(A \cap B) + P(A^c \cap B)$$

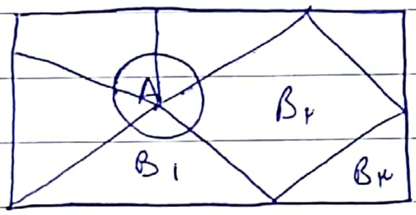
$$P(A \cap B) = P(A) \cdot P(B \mid A)$$

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

$$P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1) \cdot P(A_2 \mid A_1) \cdot P(A_3 \mid A_1 \cap A_2) \dots P(A_n \mid A_1 \cap \dots A_{n-1})$$

برای هر n رخداد که همه امکان الوقوع باشند.

فضای نمونه‌ای Ω



تجزیه احتمال کل (قانون احتمال کل):

اگر رخدادهای B_1 تا B_n یک افراز از فضای نمونه‌ای باشند.

(B_1 تا B_n دو به دو ناسازگارند و $P(B_1 \cup B_2 \cup \dots \cup B_n) = 1$
$= P(\Omega)$)

$$P(A) = \sum_{i=1}^{n} P(A \cap B_i) = \sum_{i=1}^{n} P(B_i)\, P(A \mid B_i)$$

قانون بیز:

اگر رخدادهای B_1 تا B_n یک افراز از فضای نمونه‌ای باشد به طوری که $P(B_i) \neq 0$, for $i=1:n$

آنگاه برای هر رخداد A از فضای نمونه‌ای که $P(A) \neq 0$ خواهیم داشت:

معلوم: $P(B_i)$

معلوم: $P(A \mid B_i)$

مجهول: $P(B_i \mid A)$

Subject:

$$P(B_i|A) = \frac{P(B_i \cap A)}{P(A)}$$

با استفاده از قانون احتمال کل ← $P(A) = \sum_{k=1}^{n} P(B_k) \cdot P(A|B_k)$

$$= \frac{P(B_i) \cdot P(A|B_i)}{\sum_{k=1}^{n} P(B_k) \cdot P(A|B_k)}$$

استقلال n رخداد:

n رخداد A_1 تا A_n از هم مستقل هستند (توأماً مستقل) اگر برای هر زیرمجموعه n تایی k تایی (k≥۲) از این n رخداد داشته باشیم:

$$\underbrace{P(A_{i1} \cap A_{i2} \cap \dots \cap A_{ik})}_{\text{هر زیرمجموعه k تایی}} = P(A_{i1}) \cdot P(A_{i2}) \dots P(A_{ik})$$

Pairwise

اگر n رخداد بصورت زوجی (۲به۲) از هم مستقل باشند آیا n رخداد توأماً از هم مستقل خواهند بود؟ خیر

مثال نقض: فرض کنید یک جعبه داریم که شامل ۴ بلیت است که بلیت ۴ به صورت زیر هستند

$t_1: 112$ $t_2: 121$ $t_3: 211$ $t_4: 222$

آزمایش تصادفی: انتخاب تصادفی یک بلیت از جعبه.

Subject:

A_1: یک مکان اول عدد یک بیت باشد

A_2: ـ ـ ـ ـ دوم ـ ـ ـ

A_3: ـ ـ ـ ـ سوم ـ ـ ـ

$P(A_1)=\frac{1}{2} \qquad P(A_2)=\frac{1}{2} \qquad P(A_3)=\frac{1}{2}$

$$P(A_1 \cap A_2 \cap A_3) = P(A_1) \cdot P(A_2) \cdot P(A_3)$$
$$0 \neq \frac{1}{8}$$

A_1 , A_2 , A_3 به صورت زوجی مستقل هستند که ترا با هم مستقل نیستند.

$$P(A_1) \cdot P(A_2) = P(A_1 \cap A_2) = \frac{1}{4}$$

$$P(A_1) \cdot P(A_3) = P(A_1 \cap A_3) = \frac{1}{4}$$

$$P(A_2) \cdot P(A_3) = P(A_2 \cap A_3) = \frac{1}{4}$$

اگر برای n رخداد داشته باشیم $P(A_1 \cap A_2 \cap \dots A_n) = P(A_1) \cdot P(A_2) \cdot \dots P(A_n)$

آیا می توان نتیجه گرفت رخدادها به صورت زوجی از هم مستقل هستند؟

مثال نقض: فرض کنید ۲ تاس پرتاب می کنیم رخدادهای زیر را تعریف می کنیم:

A_1: { تاس اول ۱ یا ۲ یا ۳ بیاید }

A_2: { ـ ۳ یا ۴ یا ۵ ـ ـ }

A_3: { مجموع تاس ها ۹ باشد }

Subject:

$P(A_1)=\frac{1}{2}$ $P(A_2)=\frac{1}{2}$ $P(A_3)=\frac{4}{36}$

$P(A_1\cap A_2\cap A_3) = P(A_1).P(A_2).P(A_3) = \frac{1}{36}$

$P(A_1\cap A_2)= \frac{1}{4} = \frac{1}{2}\times\frac{1}{2}$

$P(A_1\cap A_3) = \frac{1}{36}$ $\{3,6\}$

$P(A_2\cap A_3) = \frac{3}{36}$ $\{3,6\}$ $\{4,5\}$

۹۶,۱۲,۶

جلسه چهارم

احتمال شرطی :

$P(A\cap B) = P(A).P(B)$ رخداد A, B از هم مستقل هستند

$P(A\cap B|C) = P(A|C).P(B|C)$ دو رخداد A, B به شرط C از هم مستقل هستند

مثال: یک شرکت بیمه اعتقاد دارد که افراد جامعه به دو گروه بی‌احتیاط و با احتیاط تقسیم می‌شوند. افراد

$P(A|B)$ $P(A|B^c)$

با احتیاط هر سال با احتمال ۲ درصد تصادف می‌کنند و افراد بی‌احتیاط هر سال با احتمال ٤ درصد و همچنین

می‌دانیم ۷۰٪ افراد جامعه با احتیاط هستند. یک فرد که ماشین خود را بیمه کرده است با چه احتمالی

سال دوم تصادف می‌کند به شرطی که سال اول تصادف کرده باشد؟ فرض:

تصادف کردن در سال دوم و سال اول به شرط با احتیاط بودن از هم مستقل هستند.

بی‌احتیاط

Subject:

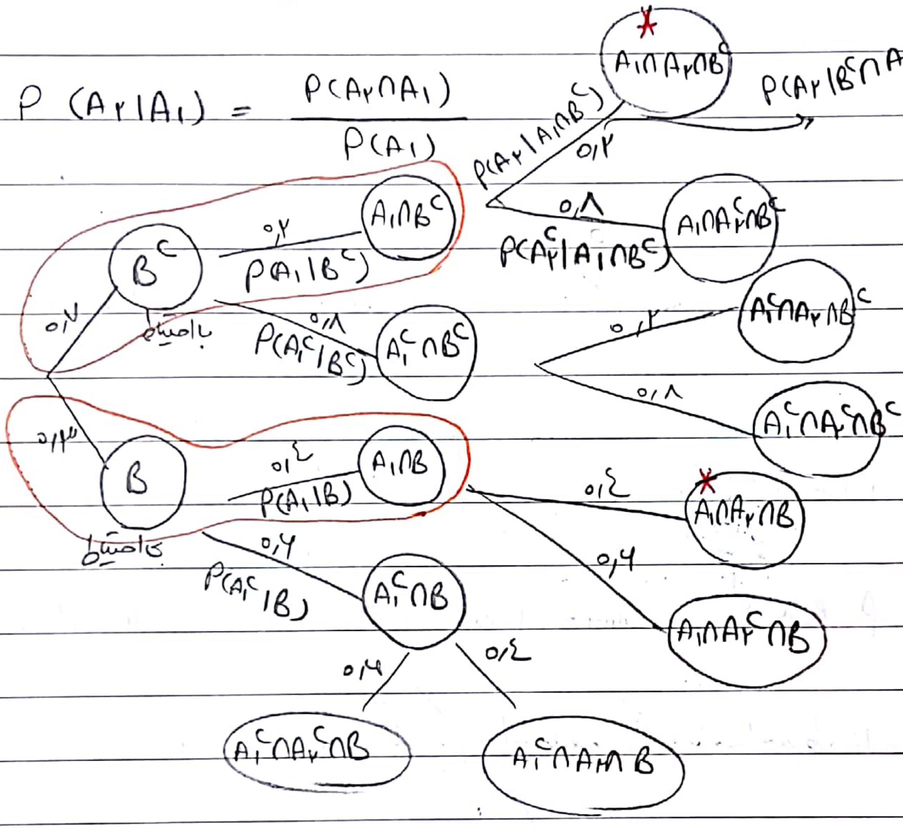
B : بی‌احتیاط بودن A_1 : تصادف در سال اول

B^c : با احتیاط بودن A_2 : تصادف در سال دوم

نکته :

$$P(A|C\cap B) = \frac{P(A\cap B\cap C)}{P(B\cap C)} = \frac{P(A\cap B|C).P(C)}{P(B|C).P(C)} = P(A|C)$$

$$P(A_2|A_1) = \frac{P(A_2\cap A_1)}{P(A_1)}$$



با توجه به نکته : $0/2$

$P(A_1\cap A_2)$: *

$$P(A_2|A_1) = \frac{P(A_2\cap A_1)}{P(A_1)}$$ احتمال شرطی

$$= \frac{P(A_1\cap A_2\cap B^c)+P(A_1\cap A_2\cap B)}{P(A_1\cap B^c)+P(A_1\cap B)} = \frac{0/7\times 0/2\times 0/2+0/3\times 0/4\times 0/4}{0/7\times 0/2+0/3\times 0/4} = 0/29$$

$P(B^c).P(A_1|B^c) \rightarrow P(B).P(A_1|B)$

$$P(B^c).P(A_1|B^c).\underbrace{P(A_2|A_1\cap B^c)} = P(B^c).P(A_1|B^c).P(A_2|B^c)$$

A_1, A_2 به شرط B^c از هم مستقل اند

Subject:

مثال: در یک مسابقه جایزه (ماشین) پشت یک در از سه در قرار داده شده است و پشت دو در دیگر بز قرار دارد. پس از آنکه مسابقه دهنده یک در را انتخاب کرد مجری یکی از دو در دیگر که پشت آن بز قرار دارد را باز می‌کند و از مسابقه دهنده می‌پرسد که آیا می‌خواهد انتخاب خود را عوض کند یا خیر. آیا تعویض انتخاب شانس برنده شدن فرد را بالا می‌برد؟

فرض کنید شخص در اول را انتخاب کند

A: ماشین پشت در اول باشد / B: مجری در دوم را باز کند.

$$P(A)=\frac{1}{3} \qquad P(A|B)=\frac{P(A\cap B)}{P(B)}=\frac{P(A)\cdot P(B|A)}{P(A\cap B)+P(A^c\cap B)}$$

$P(A)\cdot P(B|A)$ $P(A_2\cap B)+P(A_3\cap B)$

اگر بدانیم ماشین پشت در اول است چقدر احتمال دارد مجری در دوم را باز کند

$$\frac{\frac{1}{3}\times\frac{1}{2}}{\frac{1}{3}\times\frac{1}{2}+\frac{1}{3}\times 0+\frac{1}{3}\times 1}=\frac{1}{3}$$

احتمال اینکه ماشین پشت در دیگر باشد $\frac{2}{3}$ است.

راه دوم:

		انتخاب را عوض نکند	عوض کند
$\frac{1}{3}$ D_1	$\frac{1}{2}$ D_2	car	goat
	$\frac{1}{2}$ D_3	car	goat
$\frac{1}{3}$ D_2	0 D_2	—	—
	1 D_3	goat	car
$\frac{1}{3}$ D_3	1 D_2	goat	car
	0 D_3	—	—

احتمال حضور ماشین پشت در D_1

احتمال باز کردن در D_2 یا D_3 توسط مجری

$$P=\frac{\frac{1}{3}\times\frac{1}{2}+\frac{1}{3}\times\frac{1}{2}}{\frac{1}{3}\times 1+\frac{1}{3}\times 1+\frac{1}{3}\times\frac{1}{2}+\frac{1}{3}\times\frac{1}{2}}$$

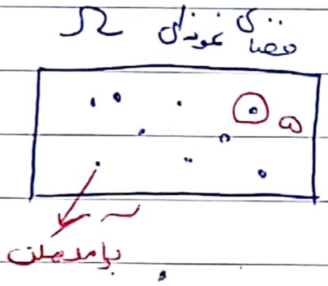
اگر تصمیم خود را عوض نکند با چه احتمالی برنده می‌شود؟

Subject: ۹۴ | ۱۲ | ۸

جلسه پنجم

فصل ۳ متغیر تصادفی

نه متغیره، نه تصادفی ← بلکه یک تابع است.



آزمایش تصادفی

انتخاب تصادفی یک دانشجو از دانشجویان کلاس

$y = X(\omega)$

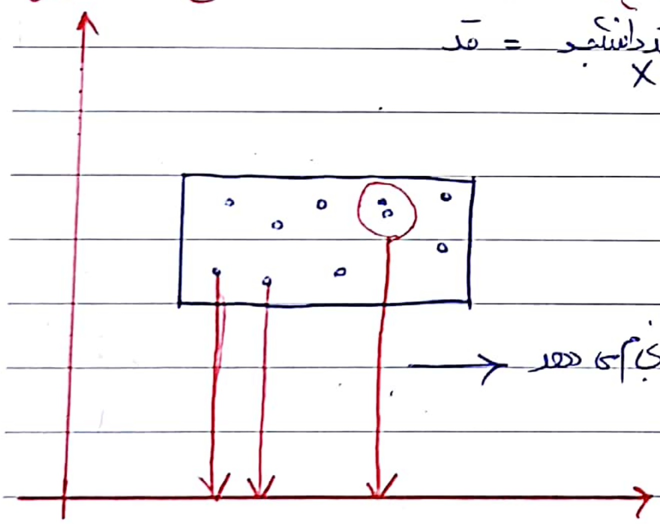
عدد حقیقی — تابع — مربوط به یک پیامد آزمایش تصادفی است

(دانشجو) وزن دانشجو = وزن / (دانشجو) قد دانشجو = قد

X

$H = \frac{X}{2.5}$

قد بر حسب اینچ



متغیر تصادفی این نگاشت را انجام می‌دهد

real حقیقی

آزمایش تصادفی: پرتاب یک سکه

فضای نمونه ← T و H

میزان جایزه: متغیر تصادفی است.

اگر H آمد ۱۰۰۰ تومان می‌برد — $1000 = X(H)$

" " T " ۵۰۰ " " — $500 = X(T)$

Subject:

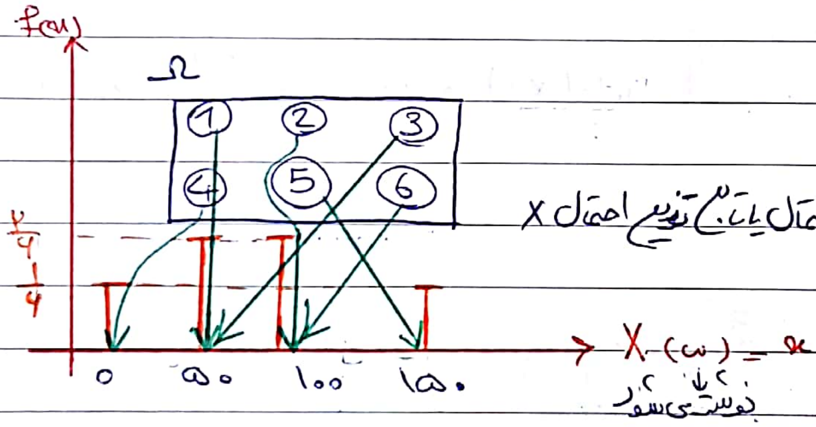
متغیر تصادفی یک تابع از فضای نمونه ای به اعداد حقیقی است.

متغیر تصادفی یک نگاشت از فضای نمونه ای به اعداد حقیقی فراهم می آورد.

مثال: آزمایش تصادفی پرتاب تاس را در نظر بگیریم

نگاشت از فضای نمونه ای به اعداد حقیقی را به صورت زیر در نظر می گیریم.

فرد تاس را به صورت تصادفی پرتاب می کند اگر تاس ۱ یا ۳ بیاید ۵۰ تومان می برد اگر تاس ۲ یا ۴ بیاید ۱۰۰ تومان برنده می شود و اگر تاس ۶ بیاید هیچ پولی برنده نمی شود و اگر تاس ۵ بیاید ۱۵۰ تومان برنده می شود.



تابع جرم احتمال یا تابع توزیع احتمال X

در حالت گسسته صحیح است { چون احتمال هر پیشامد از آزمایش تصادفی برآمده شده باشد که توسط متغیر تصادفی (تابع X) به مقدار x نگاشت می شود. $f_X(x) = P(X = x)$

میزان احتمال اینکه در آزمایش پرتاب تاس به گونه ای باشد که ۵۰ تومان جایزه برنده شویم.

$f_X(50) = \frac{2}{6}$ در مثال فوق

متغیر تصادفی X گسسته است اگر مقادیر آن گسسته باشد.

IDEA

Subject:

تابع جرم احتمال متغیر تصادفی گسسته ((Pmf) probility mass function)

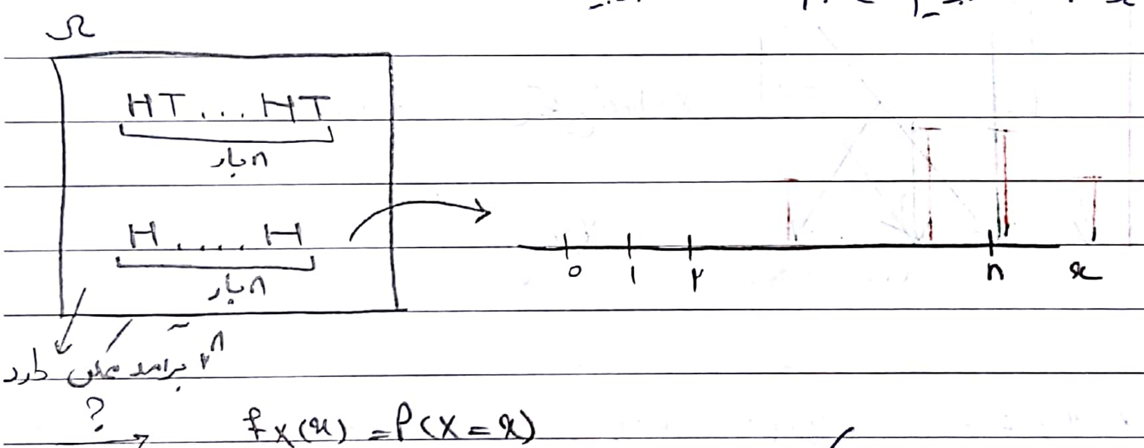
تابع توزیع احتمال (probility distribution)

- تابع $f_X(x)$ ، تابع جرم احتمال متغیر تصادفی گسسته است اگر برای هر برآمد ممکن ω از آزمایش تصادفی $X(\omega)=x$ داشته باشیم:

① $f_X(x)\geq 0$

② $\sum f_X(x)=1$, ③ $P(X=x)=f(x) \longrightarrow P(X(\omega)=x)=f(x)$

مثال: فرض کنید یک آزمایش تصادفی n بار یک سکه را پرتاب می‌کنیم متغیر تصادفی X را تعداد Hها در n پرتاب در نظر بگیریم تابع جرم احتمال X را بیابید.



$f_X(x)=P(X=x)$

$f_X(0)=P(X=0)$ یعنی احتمال اینکه در n پرتاب سکه هیچ H نیاید.

$$f_X(x)=\frac{\binom{n}{x}}{2^n} \overset{*}{=} \binom{n}{x}\cdot\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{n-x}$$

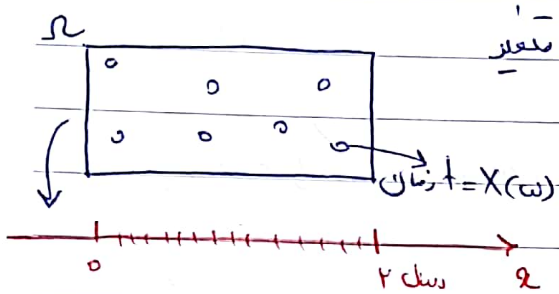
Subject:

اگر سکه غیرمنصف باشد و با احتمال p ، H را بیابید ، آنگاه تابع جرم احتمال X را بیابید.

$$\underbrace{f_X(x) = \binom{n}{x} p^x \cdot (1-p)^{n-x}}$$

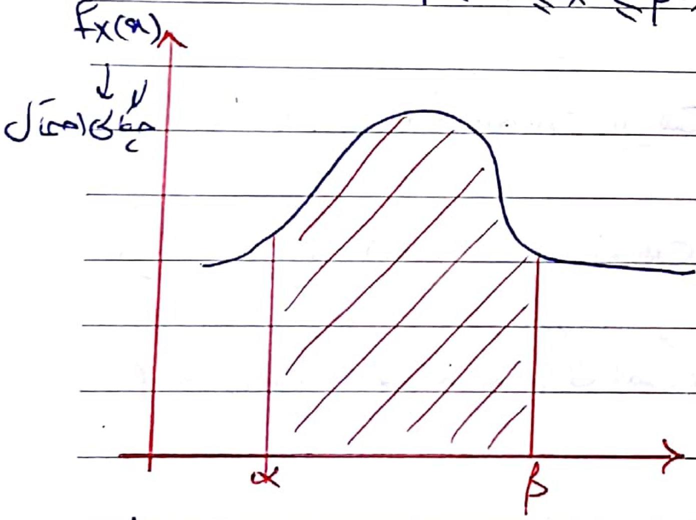
* تابع توزیع احتمال دوجمله‌ای

آزمایش تصادفی انتخاب یک کامپیوتر از ده کامپیوتر موجود متغیر



تصادفی X را طول عمر کامپیوتری که به تصادف انتخاب شده است در نظر بگیریم.

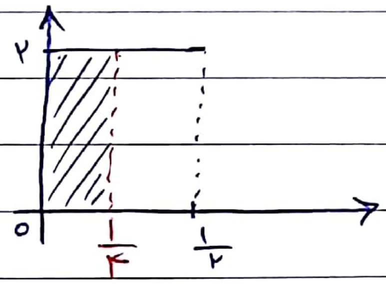
- متغیر تصادفی پیوسته متغیر تصادفی است که برد تابع X پیوسته باشد. برای متغیر تصادفی پیوسته $P(X=x)=0$ اما مطلوب است احتمال $P(\alpha \leq X \leq \beta)$



$$\int_\alpha^\beta f_X(x)\,dx = P(\alpha < X < \beta)$$

$$\int_{-\infty}^{+\infty} f_X(x)\,d(x) = 1$$

$$f_X(x) \geq 0$$



- $f(x)$ معنی احتمال نمی‌دهد و باید $f_X(x) \geq 0$

اما $\int_\alpha^\beta f_X(x)\,dx$ به معنی احتمال است در بازه (α, β)

IDEA

Subject:

$$P\left(0<x<\frac{1}{4}\right)=\frac{1}{2}=\int_{0}^{\frac{1}{4}} 2\,dx$$

تابع چگالی احتمال

$$P\left(X=\frac{1}{4}\right)=0$$

تابع چگالی احتمال (Probility density function) (pdf):

تابع توزیع احتمال

تابع توزیع احتمال برای یک متغیر تصادفی پیوسته به صورت $f_X(x)$ تعریف می‌شود که ۳ خاصیت زیر را دارد:

① $\forall x \in \text{Real} \quad f_X(x) \geq 0$

② $\int_{-\infty}^{+\infty} f_X(x)\,dx=1$ ③ $P(\alpha<X<\beta)=\int_{\alpha}^{\beta} f(x)\,dx$ یا

$$P(\alpha \leq X \leq \beta)=\int_{\alpha}^{\beta} f(x)\,dx \qquad (P(X=\alpha)=0)$$

نکته *) $f_X(x)$ در حالت پیوسته (تابع چگالی احتمال) به معنای احتمال $P(X=x)$ نیست. $P(X=x)=0$ (در حالت پیوسته) است. در $f_X(x)$ در حالت پیوسته

مفهوم احتمال در اینجا مساحت زیر آن است که به معنای احتمال دارد.

Subject:
۹۶/۱۲/۱۳

جلسه ششم

مثال : طول عمر یک کامپیوتر یک متغیر تصادفی پیوسته با تابع توزیع احتمال زیر است . احتمال

یک کامپیوتر بین ۵۰ تا ۱۵۰ روز عمر می‌کند؟

$f(x) = \alpha e^{-\alpha x}$

تابع توزیع احتمال متغیر تصادفی نمایی

x: عمر کامپیوتر بر حسب روز

$$f_X(x) = \begin{cases} \lambda e^{-\frac{x}{100}} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

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

$$\int_{-\infty}^{+\infty} f(x)\,dx = \int_0^{+\infty} \lambda e^{-\frac{x}{100}}\,dx = 1$$

$$\longrightarrow -100\lambda e^{-\frac{x}{100}} \Big|_0^{+\infty} = 1 \Longrightarrow \lambda = \frac{1}{100}$$

$$\int_{50}^{150} \frac{1}{100} e^{-\frac{x}{100}}\,dx = -e^{-\frac{x}{100}} \Big|_{50}^{150} = -e^{-\frac{150}{100}} + e^{-\frac{50}{100}}$$

چقدر احتمال دارد کامپیوتر دقیقاً ۵۰ روز عمر کند؟ ۰

تابع توزیع احتمال انباشته (تجمعی)

$P(X \leq x) = F_X(x)$ چقدر احتمال دارد متغیر تصادفی X مقداری کمتر یا مساوی x داشته باشد.

اگر متغیر تصادفی X گسسته باشد تابع توزیع (احتمال) تجمعی cumulative distribution function (CDF) آن به صورت زیر تعریف می‌شود :

IDEA

Subject:

$$F_X(x) = P(X \leq x) = \sum_{t \in -\infty}^{x} f(t) \qquad -\infty < x < +\infty$$

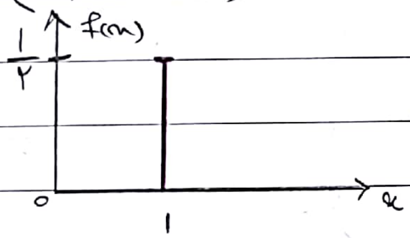
$t \leq x$:میتوان نوشت

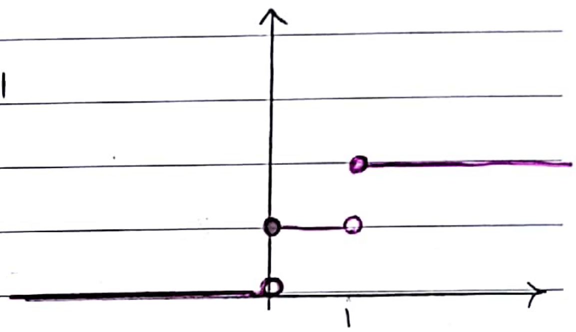
اگر تابع جرم احتمال متغیر تصادفی X به صورت زیر باشد تابع توزیع تجمعی X را بیابید.

$$f_X(x) = \begin{cases} \frac{1}{2} & x=0 \\ \frac{1}{2} & x=1 \\ 0 & o.w \end{cases} \qquad F_X(x) = \begin{cases} 0 & -\infty < x < 0 \\ \frac{1}{2} & 0 \leq x < 1 \\ 1 & 1 \leq x < +\infty \end{cases}$$

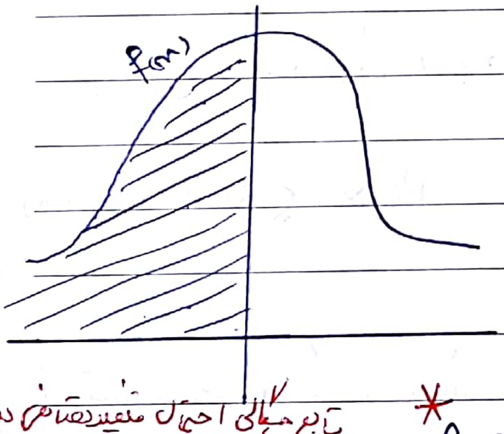
$$P(X=1) = ? = f(1) = F(1) - F(0) = F(1,1) - F(0.9) = 1 - \frac{1}{2} = \frac{1}{2}$$

$$P(-1 < X < 2) = F(2) - F(-1) = 1$$





اگر متغیر تصادفی X پیوسته باشد تابع توزیع تجمعی آن به صورت زیر تعریف می شود.



$$F_X(x) = P(X \leq x) = \int_{-\infty}^{+\infty} f(t)dt$$

$$-\infty < x < +\infty$$

← در حالت پیوسته و گسسته مقادیر منفی احتمال نداریم.

تابع چگالی احتمال متغیر تصادفی پیوسته

$$* \quad P(\alpha \leq X \leq \beta) = F_X(\beta) - F_X(\alpha)$$

$$* \quad F'_X(x) = f(x)$$

$$\int_{\alpha}^{\beta} f(x)\,dx$$

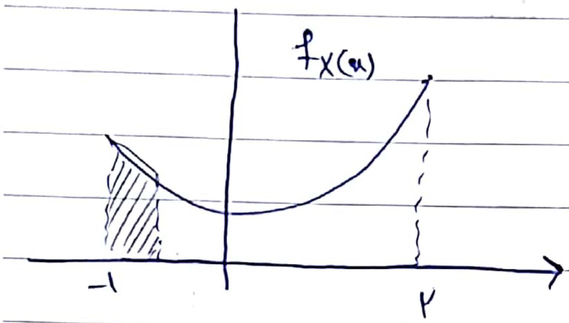
مشتق گرفتن از تابع توزیع تجمعی

Subject:

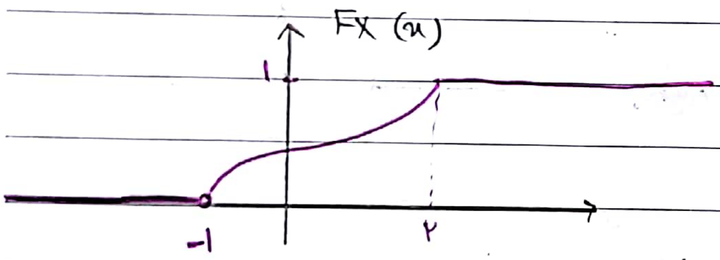
مثال: فرض کنید یک دستگاه دماسنج مایعی خطا ثبت می‌کند. متغیر تصادفی X را میزان خطای دستگاه تعریف می‌کنیم. تابع توزیع چگالی احتمال X به صورت زیر است:

$$f_X(x) = \begin{cases} \frac{x^2}{3} & -1 < x < 2 \quad \text{برحسب درجه} \\ 0 & o.w \quad \text{در سایر موارد} \end{cases}$$

الف) تابع توزیع تجمعی X را بیابید:



$$F_X(x) = \begin{cases} \int_{-\infty}^{-1} 0\, dt = 0 & -\infty < x < -1 \\ \int_{-1}^{x} \frac{t^2}{3} dt = \frac{x^3}{9} + \frac{1}{9} & -1 < x < 2 \\ 1 & x \geq 2 \end{cases}$$



ب) چقدر احتمال دارد خطای دستگاه کمتر از ۱ درجه باشد؟

$$\int_{-\infty}^{1} f(x)\,dx = F_X(1) = \frac{2}{9}$$

ج) چقدر احتمال دارد که خطای دستگاه بین ۱ تا ۲ درجه باشد؟

$$P(1 \leq X \leq 2) = F(2) - F(1) = 1 - \frac{2}{9} = \frac{7}{9}$$

Subject:

اگر یک سوزن را که محور x ها پرتاب نماییم به طوری که سوزن همراه به طور یکنواخت روی یک نقطه از بازه $[0,5]$ بیافتد. آنگاه تابع توزیع احتمال X (محل برخورد سوزن با محور x ها) را بیابید.

$$F_X(x) = P(X < x) = \frac{x-0}{5-0} = \frac{x}{5}$$

چقدر احتمال دارد محل برخورد سوزن روی نقاط $X < x$ باشد.

$0 < x < 5$

$$F'_X(x) = \frac{\partial F(x)}{\partial x} = \frac{1}{5}$$

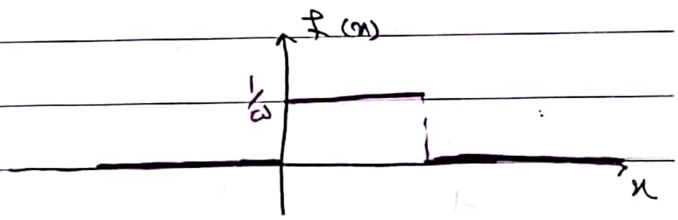
چقدر احتمال دارد سوزن روی نقاط بین ۰ تا ۲ بیافتد؟

$$\int_0^2 \frac{1}{5}\, dx = 2/5$$

چگالی

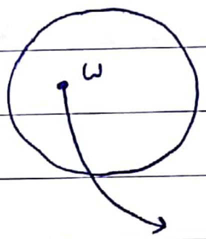
$$\int_0^\alpha f_X(x) = \frac{\alpha}{5}$$

$\frac{1}{5}$



تابع توزیع متغیر تصادفی یکنواخت

تابع توزیع توأم :



$$P(\alpha < x < \beta \;,\; \omega < y < \gamma)$$

قد $X(\omega) = 170$

وزن $y(\omega) = 80$

$$\int_\omega^\gamma \int_\alpha^\beta f(x,y)\, dx\, dy$$

IDEA

Subject:

مثال: فرض کنید یک کانال ارتباطی باینری داریم اگر X را بیت ارسالی توسط فرستنده در نظر بگیریم و Y را بیت دریافتی توسط گیرنده

$f(x,y)$	$X=0$	$X=1$	
$Y=0$	۰٫۴۵	۰٫۰۳	← ۰٫۴۸ احتمال دریافت ۰
$Y=1$	۰٫۰۵	۰٫۴۷	← ۰٫۵۲ احتمال دریافت ۱
احتمال ارسال ۰ / احتمال ارسال ۱	۱/۲	۱/۲	

$$\sum_y \sum_x f(x,y) = 1$$

توزیع حاشیه‌ای x

	$X=0$	$X=1$
$f_X(x)$	$\frac{1}{2}$	$\frac{1}{2}$

y ...

	$y=0$	$y=1$
$f_Y(y)$	۰٫۴۸	۰٫۵۲

جلسه هفتم ۹۴/۱۲/۵

تابع توزیع توأم برای متغیرهای تصادفی گسسته x و y (Joint Probality mass function)

تابع توزیع احتمال توأم متغیرهای تصادفی گسسته x و y به صورت $f_{xy}(x,y)$ تعریف می‌شود که دارای خواص زیر است:

① $f(x,y) \geq 0$

② $\sum_y \sum_x f(x,y) = 1$

③ $P[(x,y) \in A] = \sum\sum_A f(x,y)$

Subject:

مثال: درون سبد ۳ توپ آبی، ۲ توپ قرمز و ۳ توپ سبز وجود دارد. ۲ توپ به تصادف از سبد خارج می‌کنیم. متغیر تصادفی X را تعداد توپ‌های آبی و متغیر تصادفی Y را تعداد توپ‌های قرمز خارج شده از سبد تعریف می‌کنیم.

الف) تابع توزیع احتمال توأم X و Y را بدست آورید.

$f(x,y) = P(X=x \cap Y=y)$

چقدر احتمال دارد x توپ آبی و y توپ قرمز از سبد خارج شود

$f(1,1)$: ۱ توپ آبی، ۱ توپ قرمز

آبی — قرمز — سبز

$$f(x,y) = \frac{\binom{3}{x}\binom{2}{y}\binom{3}{2-x-y}}{\binom{8}{2}} \rightarrow f(x,y) = \begin{cases} \dfrac{\binom{3}{x}\binom{2}{y}\binom{3}{2-x-y}}{\binom{8}{2}} & x+y \leq 2 \\ 0 & o.w \end{cases}$$

$x+y \leq 2$ حتما باید محدوده مشخص شود.

$f(x,y)$	$X=0$	$X=1$	$X=2$	$f_Y(y)$
$Y=0$	$\frac{3}{28}$	$\frac{9}{28}$	$\frac{3}{28}$	$\frac{15}{28}$
$Y=1$	$\frac{3}{14}$	$\frac{3}{14}$	0	$\frac{12}{28}$
$Y=2$	$\frac{1}{28}$	0	0	$\frac{1}{28}$
$f_X(x)$	$\frac{10}{28}$	$\frac{15}{28}$	$\frac{3}{28}$	

توزیع حاشیه‌ای Y

*) از روی تابع توزیع احتمال توأم می‌توان تابع توزیع هر کدام از متغیرها را بدست آورد.

توزیع حاشیه‌ای X

ب) چقدر احتمال دارد $x+y \leq 1$ باشد؟

$$P(x+y \leq 1) = f(0,0) + f(0,1) + f(1,0) = \frac{3}{28} + \frac{9}{28} + \frac{3}{14} = \frac{18}{28}$$

Subject:

تابع توزیع احتمال توأم برای متغیرهای تصادفی پیوسته X و Y به صورت $f(x,y)$ تعریف می شود

که دارای خواص زیر است: ① $f(x,y) \geq 0$ ، ② $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\,dx\,dy = 1$

③ $P[(X,Y) \in A] = \iint_A f(x,y)\,dx\,dy \Longrightarrow$

برای هر ناحیه از صفحه x و y

$$P[\alpha < X < \beta ,\ \omega < Y < \gamma] = \int_{\alpha}^{\beta}\int_{\omega}^{\gamma} f(x,y)\,dy\,dx$$

مثال: یک سایت خدمت رسانی دو نوع خدمت ارائه می دهد. متغیر تصادفی X را نسبتی از زمان در نظر می گیریم که از سایت خدمت ① درخواست می شود. و متغیر تصادفی Y را نسبتی از زمان در نظر می گیریم که از سایت خدمت ② درخواست می شود. اگر تابع توزیع احتمال توأم X ، Y به صورت زیر باشد، احتمال $P[0 < X < \frac{1}{2} ,\ \frac{1}{4} < Y < \frac{1}{2}]$ را بدست آورید

(دقیقا احتمال اینکه نسبتی از زمان که خدمت ① درخواست می شود بین 0 تا $\frac{1}{2}$ و نسبتی از زمان که خدمت ② درخواست می شود بین $\frac{1}{4}$ تا $\frac{1}{2}$ باشد)

$$f(x,y) = \begin{cases} \frac{2}{5}(2x+3y) & 0 \leq x \leq 1 \quad 0 \leq y \leq 1 \\ 0 & O.W \end{cases}$$

Subject:

$$P\left[0 \leq X \leq \frac{1}{2} \cap \frac{1}{4} \leq Y \leq \frac{1}{2}\right] = \int_{\frac{1}{4}}^{\frac{1}{2}} \int_{0}^{\frac{1}{2}} \frac{2}{5}(2x+3y)\, dx\, dy$$

$$= \int_{\frac{1}{4}}^{\frac{1}{2}} \left(\frac{2x^2}{5} + \frac{6xy}{5}\right)\Bigg|_{x=0}^{x=\frac{1}{2}} =$$

$$\left(\frac{y}{10} + \frac{3y^2}{10}\right)\Bigg|_{\frac{1}{4}}^{\frac{1}{2}} = \frac{1}{10}\left[\left(\frac{1}{2} + \frac{3}{4}\right) - \left(\frac{1}{4} + \frac{3}{16}\right)\right] = \frac{13}{160}$$

تابع توزیع حاشیه‌ای متغیر تصادفی X و Y به صورت زیر براساس $f(x,y)$ تعریف می‌شود:

$$f_X(x) = g_X(x) = \sum_{y} f(x,y)$$

$$f_Y(y) = h_Y(y) = \sum_{x} f(x,y)$$

برای حالت گسسته ✓

$$f_X(x) = g_X(x) = \int_{-\infty}^{+\infty} f(x,y)\, dy$$

$$f_Y(y) = h_Y(y) = \int_{-\infty}^{+\infty} f(x,y)\, dx$$

برای حالت پیوسته

در مثال قبل احتمال $P\left[\frac{1}{4} < Y < \frac{1}{2} \,\middle|\, 0 < X < \frac{1}{2}\right]$ ؟

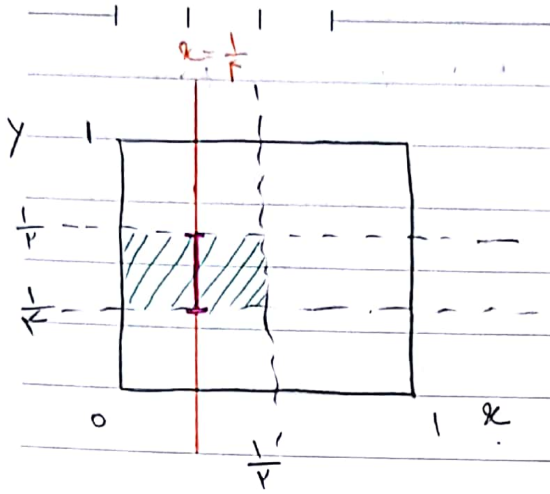
$$P\left[\frac{1}{4} < Y < \frac{1}{2} \,\middle|\, 0 < X < \frac{1}{2}\right] = \frac{P\left[\frac{1}{4} < Y < \frac{1}{2} \cap 0 < X < \frac{1}{2}\right]}{P\left(0 < X < \frac{1}{2}\right)}$$

$$= \frac{\int_{\frac{1}{4}}^{\frac{1}{2}} \int_{0}^{\frac{1}{2}} f(x,y)\, dx\, dy}{\int_{0}^{\frac{1}{2}} f_X(x)\, dx} = \frac{\frac{13}{160}}{\frac{2}{5}}$$

$$\int_{0}^{\frac{1}{2}} f_X(x)\, dx = \frac{2x^2}{5} + \frac{3x}{5}\Bigg|_{0}^{\frac{1}{2}} = \frac{2}{5}$$

کل y $\int_{0}^{1} \int_{0}^{\frac{1}{2}} f(x,y)\, dx\, dy$

Subject:



احتمال $P[\frac{1}{4}<Y<\frac{1}{2} \mid X=\frac{1}{4}]$

این احتمال وجود دارد و قابل محاسبه است

$$P[\alpha<X<\beta \mid Y=y] = \int_{\alpha}^{\beta} f(x\mid y)\,dx$$

برای متغیرهای تصادفی پیوسته X و Y

تابع توزیع احتمال شرطی

$P[-\infty<X<x \mid Y=y_1]$

$$f(x\mid y)=\frac{f(x,y)}{f_y(y)}$$

اثبات: (زمانی که نسبت به y بیابیم برای فهم بیشتر است)

$$P(-\infty<X<x \mid y_1<Y<y_1+\Delta y) \underset{\Delta y\to 0}{\overset{\lim}{=}} \frac{\int_{-\infty}^{x}\int_{y_1}^{y_1+\Delta y} f(x,y)\,dy\,dx}{\int_{y_1}^{y_1+\Delta y} g_y(y)\,dy}$$

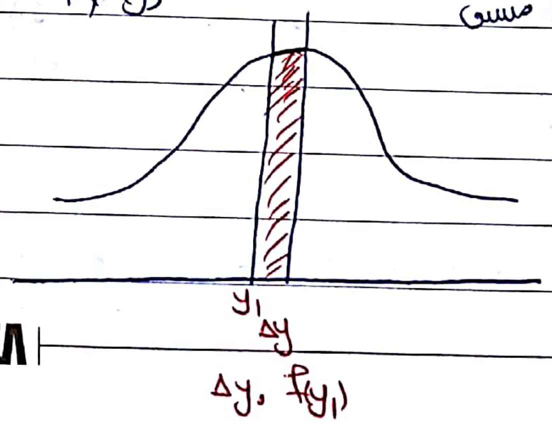
$F_X(x \mid y_1<Y<y_1+\Delta y)$

$= F_X(x\mid y_1)$

$$\frac{\Delta y \int_{-\infty}^{x} f(x,y_1)\,dx}{\Delta y\, g_y(y_1)} = \frac{\int_{-\infty}^{x} f(x,y_1)\,dx}{f_y(y)}$$

نسبت به x مشتق

$$f(x\mid y_1)=\frac{f(x,y_1)}{f_y(y_1)}$$



Subject:
۹۴ | ۱۲ | ۲۰

جلسه هشتم

تابع توزیع احتمال شرطی:

یادآوری:

$f(x,y) \longleftarrow X,Y$

$$P(\alpha < X < \beta \mid Y=y) = \frac{\overbrace{P(\alpha<X<\beta \cap Y=y_1)}^{0}}{\underbrace{P(Y=y_1)}_{0}}$$

$P(Y=y_1=0)$

$$= \int_\alpha^\beta A(x,y_1)\,dx = \int_\alpha^\beta \frac{f(x,y_1)}{f_Y(y_1)}\,dx$$

تعریف: اگر X و Y دو متغیر تصادفی گسسته یا پیوسته باشند تابع توزیع احتمال شرطی X به شرط $Y=y$

به صورت زیر تعریف می‌شود:

$$f(x|y) = \frac{f(x,y)}{h(y)} = \frac{f(x,y)}{f_Y(y)}, \quad h(y)>0$$

همچنین به طور مشابه تابع توزیع احتمال Y به شرط $X=x$ به صورت زیر است:

$$f(y|x) = \frac{f(x,y)}{g(x)} = \frac{f(x,y)}{f_X(x)} \quad f_X(x)>0$$

*) اگر $g(x)=0$ باشد $f(y|x)=0$

در مثال سبد توپ تابع توزیع احتمال توأم X و Y به صورت زیر بود. (۳ آبی، ۲ قرمز، ۳ سبز)

$P(X=0 \mid Y=1)$ چقدر است؟

X: تعداد توپ‌ها که آبی خارج شده

Y: تعداد توپ‌های قرمز خارج شده

$f(x,y)$	$X=0$	$X=1$	$X=2$	$f_Y(y)$
$Y=0$	$\frac{3}{28}$	$\frac{9}{28}$	$\frac{3}{28}$	$\frac{15}{28}$
$Y=1$	$\frac{3}{14}$	$\frac{3}{14}$	0	$\frac{6}{14}$
$Y=2$	$\frac{1}{28}$	0	0	$\frac{1}{28}$

$$\frac{P(X=0 \cap Y=1)}{P(Y=1)} = \frac{\frac{3}{14}}{\frac{6}{14}} = \frac{1}{2}$$

$$P(X=0 \mid Y=1) = f_{X|Y}(0 \mid 1) = \frac{f(x,y)}{f_Y(y)} = \frac{f(0,1)}{f_Y(1)} = \frac{\frac{3}{14}}{\frac{6}{14}}$$

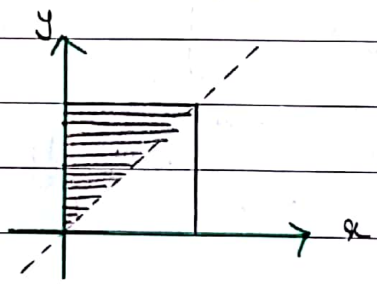
	$X=0$	$X=1$	$X=2$
$f(x \mid y=1)$	$\frac{1}{2}$	$\frac{f(1,1)}{f_Y(1)} = \frac{\frac{3}{14}}{\frac{6}{14}}$	$\frac{f(2,1)}{f_Y(1)} = \frac{0}{\frac{6}{14}}$

مثال: تابع توزیع احتمال توأم دو متغیر تصادفی X و Y به صورت زیر است:

الف) تابع توزیع احتمال $f(y \mid x)$ را به دست آورید.

ب) $P(Y > \frac{1}{2} \mid X = 0.25)$ ← $\int_{\frac{1}{2}}^{1} f(y \mid x=0.25)$

$$f(x,y) = \begin{cases} 10xy^2 & 0<x<y<1 \\ 0 & o.w \end{cases}$$



الف)

$$f(y \mid x) = \frac{f(x,y)}{f_X(x)}$$

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)\,dy = \int_{-\infty}^{x} 0\,dy + \int_{x}^{1} 10xy^2\,dy + \int_{1}^{\infty} 0\,dy =$$

$$= \frac{10}{3} x y^3 \Big|_{y=x}^{y=1} = \frac{10}{3} x(1-x^3) \qquad 0<x<1$$

به محدوده انتگرال گیری دقت کنید

Subject:

$$f(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{15xy^2}{\frac{15}{3}x(1-x^3)} = \frac{3y^2}{(1-x^3)} \qquad 0<x<y<1$$

ب)

$$\int_{\frac{1}{2}}^{1} \frac{3y^2}{(1-(0{,}25)^3)}\,dy = \frac{8}{9}$$

* اگر $P(Y>\frac{1}{5} \mid x=0{,}25)$ بود $\longleftarrow$ $\int_{0{,}25}^{1} \frac{3y^2}{1-(0{,}25)^3}\,dy$

نکته * $P(Y>\frac{1}{2} \mid \frac{1}{4}<X<1) = \frac{P(Y>\frac{1}{2} \cap \frac{1}{4}<X<1)}{P(\frac{1}{4}<X<1)} =$

$$\frac{\int_{\frac{1}{2}}^{+\infty}\int_{\frac{1}{4}}^{1} f(x,y)\,dx\,dy}{\int_{\frac{1}{4}}^{1} f_X(x)\,dx}$$

اگر $f(x|y)$ به y وابسته نباشد، آنگاه $f(x|y) = g(x) = f_X(x)$ و

$$: \quad f(x,y) = g(x)\cdot h(y) = f_X(x)\cdot f_Y(y)$$

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

* معادل برای است. $f(x|y) = \frac{f(x,y)}{h(y)} \longrightarrow f(x,y) = h(y)\cdot f(x|y)$

$$f_X(x) = g(x) = \int_{-\infty}^{+\infty} f(x,y)\,dy = \int_{-\infty}^{+\infty} h(y)\cdot f(x|y)\,dy$$

$$= f(x|y)\int_{-\infty}^{+\infty} h(y)\,dy = f(x|y)$$

(انتگرال برابر ۱)

$\downarrow$ $f(x|y)$ را می‌توانیم بیرون بیاوریم چون $f(x|y)$ به y وابسته نیست.

Subject:

مثال: تابع توزیع احتمال توأم X و Y به صورت زیر است. $f_X(x)$ و $f_{X|Y}(x|y)$ را

بدست آورید:

$$f(x,y) = \begin{cases} \dfrac{x(1+3y^2)}{4} & 0<x<2 \ , \ 0<y<1 \\ 0 & \text{O.W} \end{cases}$$

$$f_X(x) = \int_0^1 \frac{x(1+3y^2)}{4}\,dy = \left(\frac{xy}{4} + \frac{xy^3}{4}\right)\Bigg|_{y=0}^{y=1} = \frac{x}{2}$$

$$f_Y(y) = \int_0^2 \frac{x(1+3y^2)}{4}\,dx = \left(\frac{x^2}{8} + \frac{3x^2y^2}{8}\right)\Bigg|_{x=0}^{x=2} = \frac{1+3y^2}{2}$$

$$f(x|y) = \frac{f(x,y)}{f_Y(y)} = \frac{\frac{x(1+3y^2)}{4}}{\frac{1+3y^2}{2}} = \frac{x}{2} = f_X(x)$$

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{\frac{x(1+3y^2)}{4}}{\frac{x}{2}} = \frac{1+3y^2}{2} = f_Y(y)$$

$f(x,y) = f_X(x) \cdot f_Y(y)$

$f(x|y) = f_X(x)$

$f(y|x) = f_Y(y)$

$\leftarrow$ $\boxed{X \text{ و } Y \text{ از هم مستقل اند}}$

تابع توزیع احتمال توأم برابر ضرب تابع توزیع های X و Y است.

متغیرهای تصادفی مستقل (گسسته یا پیوسته):

X و Y دو متغیر تصادفی مستقل هستند اگر تابع توزیع احتمال توأم آنها برابر ضرب توزیع احتمال

حاشیه ای آنها باشد:

$$f(x,y) = f_X(x) \cdot f_Y(y) = g(x) \cdot h(y)$$

(به ازای همه x و y ها)

Subject:

با توجه به مثال صفحه ۳۱ آیا X و Y مستقل هستند؟

* یک مثال نقض کافی است برای رد استقلال

$$\underbrace{f(1,2)}_{0} \neq \underbrace{f_X(1)}_{\frac{15}{28}} \cdot \underbrace{f_Y(2)}_{\frac{1}{28}}$$

n متغیر تصادفی مستقل (گسسته یا پیوسته):

n متغیر تصادفی X_1 و X_2 و ... و X_n تواماً از هم مستقل هستند اگر و فقط اگر

$$f(x_1, x_2, \dots, x_n) = f_{X_1}(x_1) \cdot f_{X_2}(x_2) \dots f_{X_n}(x_n)$$

$$* \quad f(x_1, x_2, \dots, x_{n-1}) = \int_{-\infty}^{+\infty} f_{X_1}(x_1) f_{X_2}(x_2) \dots f_{X_{n-1}}(x_n)\, dx_{n-1}$$

$$= f_{X_1}(x_1) \dots f_{X_{n-1}}(x_{n-1}) = \underbrace{\int_{-\infty}^{+\infty} f_{X_n}(x_n)\, dx_n}_{\to 1}$$

* اگر ۲ متغیر تصادفی مستقل باشند به راحتی می‌توان تابع توزیع احتمال توام آنها ($f(x,y)$) را از ضرب تابع توزیع‌های آنها به دست آورد.

IDEA

Subject:

۹۶/۱۲/۲۲

جلسه ۳۲

مثال: طول عمر یک نوع قطعه الکترونیکی از توزیع زیر پیروی می‌کند اگر ۳ قطعه از این نوع به تصادف انتخاب می‌کنیم چه احتمالی که قطعه اول کمتر از ۲ سال، قطعه دوم بین ۱ تا ۳ سال و قطعه سوم بیش از ۲ سال عمر می‌کند.

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

x برحسب سال

X_1: طول عمر قطعه اول

X_2: طول عمر قطعه دوم

X_3: طول عمر قطعه سوم

$$\int_2^{+\infty}\int_1^{3}\int_0^{2} \underbrace{f(x_1, x_2, x_3)}\, dx_1\, dx_2\, dx_3$$

$f(x_1) \cdot f(x_2) \cdot f(x_3)$ چون مستقل هستند

$$= \left(\int_0^2 e^{-x_1}\, dx_1\right) \cdot \left(\int_1^3 e^{-x_2}\, dx_2\right) \cdot \left(\int_2^{+\infty} e^{-x_3}\, dx_3\right)$$

$= 0/037$

$f(x_1, x_2, x_3, x_4, x_5)$

$X_1 \quad X_2 \quad X_3 \quad X_4 \quad X_5$

*) $f(X_1, X_2 \mid X_3, X_4)$

اگر پیوسته باشند یعنی احتمال اشتراک X_1 و X_2 به شرط اشتراک X_3 و X_4

$$= \frac{f_{X_1, X_2, X_3, X_4}(x_1, x_2, x_3, x_4)}{f_{X_3, X_4}(x_3, x_4)}$$

Subject:

تابع توزیع حاشیه‌ای X_1 و X_2

$$f(x_1, x_2) = P(X_1 = x_1 \cap X_2 = x_2)$$

بیشتر باشد — قد — وزن

$$P(\alpha < X_1 < \beta \cap \omega < X_2 < \gamma) = \int_{\omega}^{\gamma}\int_{\alpha}^{\beta} f(x_1, x_2)\,dx_1\,dx_2$$

$$f(x_1, x_2) = \sum_{x_3}\sum_{x_4}\sum_{x_5} f(x_1, x_2, x_3, x_4, x_5)$$

$$P(\alpha < X < \beta \cap \gamma < Y < \omega ,\ r < Z < S)$$

$$= \int_{\alpha}^{\beta} f(x)\,dx \cdot \int_{\gamma}^{\omega} f(y)\,dy \cdot \int_{r}^{S} f(z)\,dz$$

*) می‌توان تابع توزیع توأم n متغیر تصادفی را به صورت زیر تعریف کرد

$$f(x_1, x_2, \ldots, x_n)$$

$$\begin{cases} \sum_{x_1}\sum_{x_2} \cdots \sum_{x_n} f(x_1, x_2, \ldots, x_n) = 1 & \text{گسسته} \\ \int_{x_1}\int_{x_2} \cdots \int_{x_n} f(x_1, \ldots, x_n) = 1 \end{cases}$$

می‌توان توزیع حاشیه‌ای هر ترکیب از n متغیر تصادفی را از آن بدست آورد به طور مثال:

$$f_{X_1}(x_1) = \sum_{x_2}\sum_{x_3} \cdots \sum_{x_n} f(x_1, \ldots, x_n)$$

$$f_{X_1}(x_1) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(x_1, \ldots, x_n)\,dx_2\,dx_3 \ldots dx_n$$

Subject:

$$f(x_1, x_2) = \sum_{x_3}\sum_{x_4}\cdots\sum_{x_n} f(x_1, x_2, \dots, x_n)$$ یا

$$f(x_1, x_2) = \int_{x_3}\int_{x_4}\cdots\int_{x_n} f(x_1, \dots, x_n)\, dx_3\, dx_4, \dots, dx_n$$

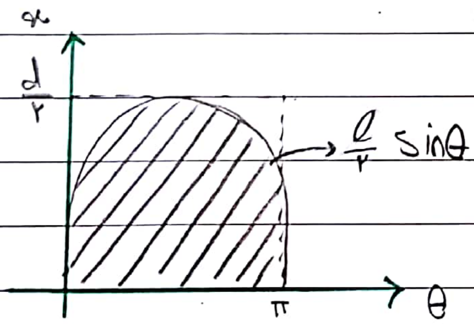
همچنین توزیع احتمالات شرطی نیز قابل محاسبه است

$$f(x_1, x_2, x_3 \mid x_4, x_5, \dots, x_n) = \frac{f_{X_1, \dots X_n}(x_1, x_2, \dots, x_n)}{f_{X_4, \dots, X_n}(x_4, \dots, x_n)}$$

مثال ۵: مسئله سوزن بوفون را در نظر بگیرید:

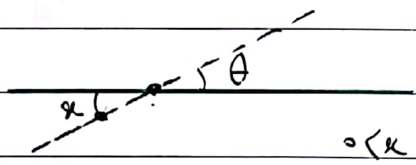
سوزنی به طول l روی یک صفحه خط کشی شده ، فاصله خطوط آن d ، $l<d$ است می اندازیم

چقدر احتمال دارد سوزن خط را قطع کند. (فرض کنید همه مکان ها و زاویه ها هم احتمال باشند)



x: فاصله مرکز سوزن تا نزدیک ترین خط

θ: زاویه سوزن با خط



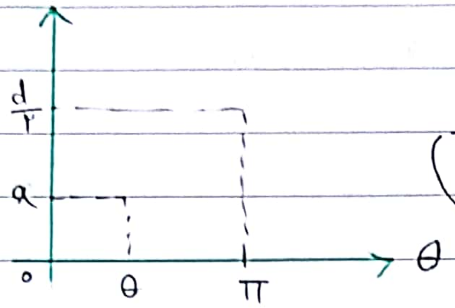
$$0 < x < \frac{l}{2}\sin\theta$$

$$\int_0^{\pi}\int_0^{\frac{l}{2}\sin\theta} f(x,\theta)\, dx\, d\theta = $$ احتمال برخورد سوزن با خط

$$\int_0^{\pi}\int_0^{\frac{d}{2}} f(x,\theta)\, dx\, d\theta = 1$$

Subject:

تابع توزیع تجمعی $F(x,\theta) = P(X<x \cap \Theta<\theta) = \dfrac{x\theta}{\pi \frac{d}{2}}$



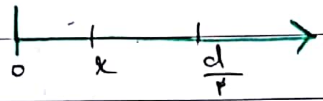
$$\frac{\left(\dfrac{\partial F(x,\theta)}{\partial\theta}\right)}{\partial x} = \frac{1}{\frac{\pi d}{2}} = \frac{2}{\pi d}$$

یکبار نسبت به θ یکبار نسبت به x

$$\Rightarrow f(x,\theta) = \underbrace{f_X(x)}_{\frac{2}{d}} \cdot \underbrace{f_\Theta(\theta)}_{\frac{1}{\pi}}$$

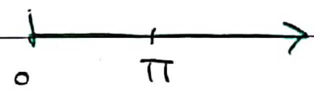
$\frac{2}{d} \times \frac{1}{\pi}$

= چون مستقل هستند



$F(x) = P(X<x) = \dfrac{x}{d/2}$

$f_X(x) = F'(x) = \dfrac{2}{d}$



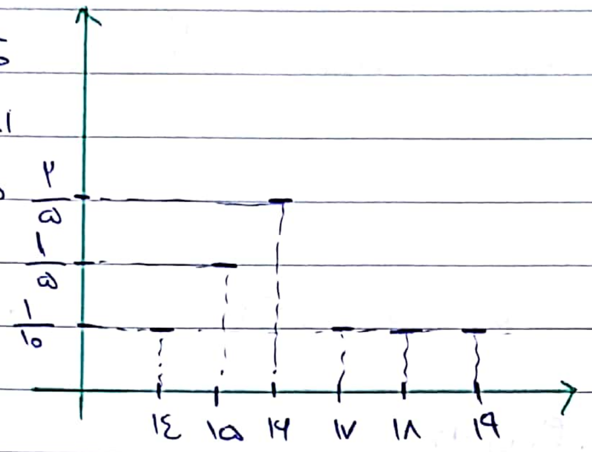
$F(\theta) = P(\Theta<\theta) = \dfrac{\theta}{\pi}$ $\quad F'(\theta) = \dfrac{1}{\pi}$

جواب $\longrightarrow \displaystyle\int_0^{\pi}\int_0^{\frac{l}{2}\sin\theta} \frac{2}{\pi d}\, dx\, d\theta = \frac{2l}{\pi d}$

Subject:

فصل چهارم

تابع توزیع جرم احتمال معدل دانشجویان



یک نفر را به تصادف انتخاب می‌کنیم ⟵ معدل : X ⟵ متغیر تصادفی

این آزمایش را ۱۰۰۰ بار تکرار می‌کنیم. میانگین معدل این هزار نفر چقدر است:

$$\text{میانگین معدل ۱۰۰۰ نفر} = \frac{\frac{1}{10}\times 14+\frac{1}{5}\times 15+\frac{2}{5}\times 16+\frac{1}{10}\times 17+\frac{1}{10}\times 18+\frac{1}{10}\times 19}{\frac{1}{10}+\frac{1}{5}+\frac{2}{5}+\frac{1}{10}+\frac{1}{10}+\frac{1}{10}}$$

↓

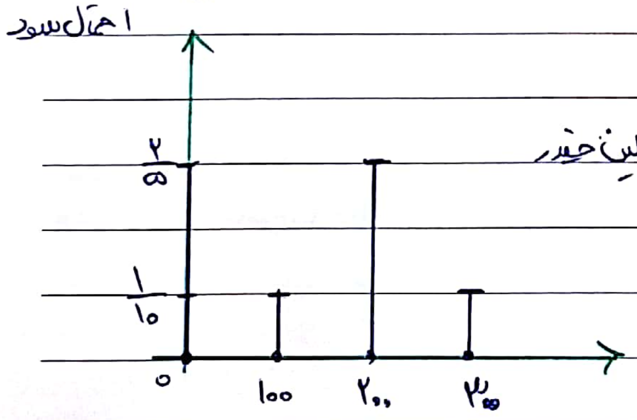
میانگین یک متغیر تصادفی

⟵ امید ریاضی : expectation

احتمالات نقش فراوانی نسبی دارند : از هر ۱۰۰۰ دانشجو ۱۰۰ نفر معدل ۱۴ دارند $= \frac{1}{10}$

یک فعالیت اقتصادی

اگر ۱۰ بار این فعالیت اقتصادی را انجام شود به طور میانگین چقدر سود می‌کنیم؟ $\frac{2}{5}\times 0+\frac{1}{10}\times 100+\frac{2}{5}\times 200+\frac{1}{10}\times 300$



Subject:

امید ریاضی متغیر تصادفی:

امید ریاضی (یا میانگین) متغیر تصادفی X در حالت گسسته به صورت زیر تعریف می‌شود:

$$E[X] = \sum_{x} x \cdot f_X(x)$$

امید ریاضی (یا میانگین) متغیر تصادفی در حالت پیوسته به صورت زیر تعریف می‌شود:

$$E[X] = \int_{-\infty}^{+\infty} x \cdot f(x)\, dx$$

مثال: یک مسئول کنترل کیفیت از بین ۷ قطعه الکترونیکی که ۴ تای آن سالم و ۳ تای آن خراب است ۳ قطعه به تصادف انتخاب می‌کند. (آزمایش تصادفی) امید ریاضی تعداد قطعات سالم موجود در نمونه ۳ تایی را بدست آورید؟

X: تعداد قطعات سالم موجود در نمونه

تعداد قطعات سالم

$$f(x) = \frac{\binom{4}{x}\binom{3}{3-x}}{\binom{7}{3}} \qquad x = 0,1,2,3$$

$$E[X] = \sum_{x=0}^{3} x \cdot f(x) = 0 \times \frac{\binom{4}{0}\binom{3}{3}}{\binom{7}{3}} + 1 \times \frac{\binom{4}{1}\binom{3}{2}}{\binom{7}{3}} + 2 \times \frac{\binom{4}{2}\binom{3}{1}}{\binom{7}{3}} + 3 \times \frac{\binom{4}{3}\binom{3}{0}}{\binom{7}{3}}$$

$$= 1/7$$

یعنی اگر این آزمایش بارها تکرار شود به طور میانگین ۱/۷ قطعه سالم در نمونه وجود خواهد داشت.

Subject:

مثال: متغیر تصادفی X طول عمر یک قطعه الکترونیکی است که تابع توزیع آن به صورت روبه‌روست

به طور میانگین طول عمر این نوع از قطعات چقدر است؟

$$f(x) = \begin{cases} \dfrac{200000}{x^3} & x > 100 \\ 0 & o.w \end{cases}$$

برحسب ساعت

$$E[X] = \int_{100}^{\infty} x \cdot \frac{200000}{x^3}\, dx = 200$$

IDEA

Subject:

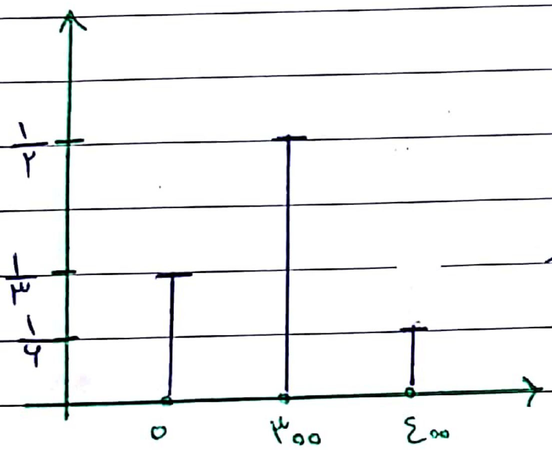
۱۳۹۷/۱۱/۲۴

جلسه دهم

مروری بر امید ریاضی:

$$E[x] = \sum_{x \in X} x \cdot f(x)$$

$$E[x] = \int_{-\infty}^{+\infty} x \cdot f(x)$$



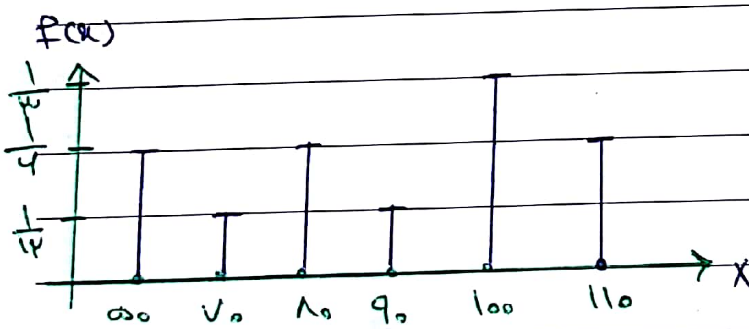
امید ریاضی امتیاز این بازی ۲۱۶ است

اگر این بازی به تعداد زیاد تکرار شود و میانگین امتیازاتی که

در این تکرارها بدست می‌آید نزدیک ۲۱۶ است.

$$E[X] = 0 \times \frac{1}{3} + 300 \times \frac{1}{2} + 400 \times \frac{1}{4}$$

تعبیر اساسی امید ریاضی:



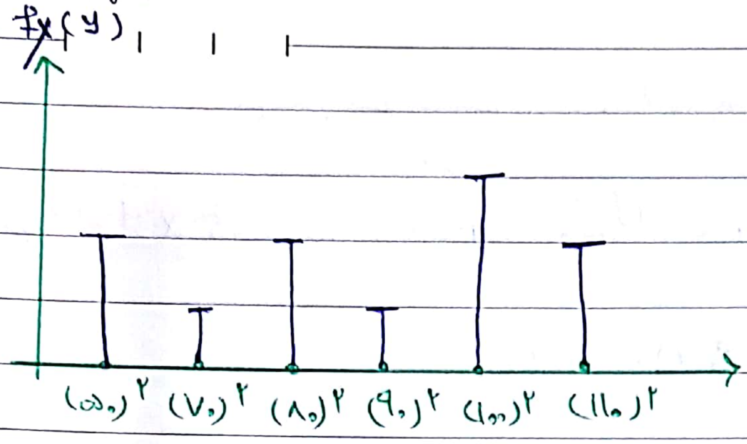
$$\underset{\text{وزن}}{E[X]} = \frac{1}{4} \times 50 + \frac{1}{12} \times 70 + \dots + \frac{1}{4} \times 110$$

$$BMI = \frac{X}{(1.7)^2} = \frac{1}{(1.7)^2} E[X]$$

$$E[\underset{y}{X^2}] = \frac{1}{4} \times (50)^2 + \frac{1}{12} \times (70)^2 + \dots + \frac{1}{4} \times (110)^2$$

$$\Rightarrow E[Y] = \sum_{y \in X} y \, f_y(y)$$

Subject:



$X \xrightarrow{g} Y = g(x) = x^2$

$0 < x < 100$

$f_Y(y) = f_X(\sqrt{y})$

$f_Y(2500) = f_X(50) = \frac{1}{6}$

$X \longrightarrow Y = g(x)$

$f_X(x) \longrightarrow f_Y(y)$

$$E[g(x)] = E[Y] = \sum_{x \in X} \underbrace{x^2}_{g(x)} f_X(x) = \sum_{x \in X} g(x) f_X(x)$$

امید ریاضی تابعی از یک متغیر تصادفی

x	-1	0	1	2
$f(x)$	$f(-1)$	$f(0)$	$f(1)$	$f(2)$

$E[X] = -1 \times f(-1) + 0 \times f(0) + 1 \times f(1) + 2 \times f(2)$

$Y = g(x) = x^2$

$$E[g(x)] = E[Y] = \sum_{y \in Y} y \cdot f_X(y)$$

y	0	1	4
$f_Y(y)$	$f_X(0)$	$f_X(1) + f_X(-1)$	$f_X(2)$

$= 0 \times f_X(0) + 1 \times (f_X(1) + f_X(-1)) + 4 \times f_X(2)$

$= 0 \times f_X(0) + 1 \times f_X(1) + f_X(-1) + 4 \times f_X(2) = \sum_{x \in X} x^2 f_X(x)$

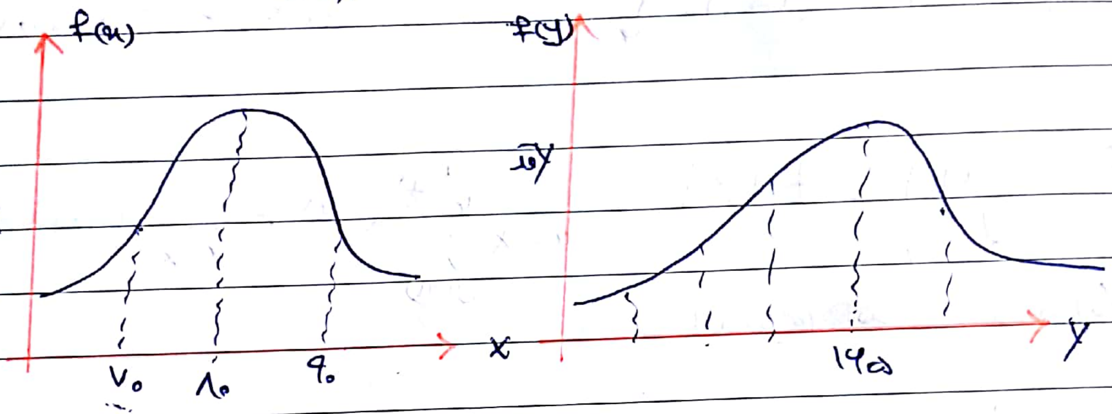
Subject:

با در اختیار داشتن تابع توزیع متغیر تصادفی X می توان امید ریاضی متغیر تصادفی $g \sim y = g(x)$

تابعی دلخواه از X را به صورت زیر بدست آورد و نیازی به محاسبه تابع توزیع متغیر y نیست.

$$E[g(x)] = \sum_{x \in X} g(x) \cdot f_X(x)$$

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



(X و Y مستقل نیستند)

$P(70 < X < 80 \cap 160 < Y < 170)$

$f(x,y)$

xy $\qquad E[X+Y]$

$$E[BMI] = E\left[\frac{X}{\left(\frac{Y}{100}\right)^2}\right] \Rightarrow \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \frac{x}{\left(\frac{y}{100}\right)^2} f(x,y)\,dx\,dy$$

با هر دو متغیر تصادفی X و Y به صورت ما اختیار داشتن $f_{XY}(x,y)$ امید ریاضی هر تابع

دلخواه $g(x,y)$ از X و Y به صورت زیر تعریف می شود.

Subject:

$$E[g(x,y)] = \sum_{y\in Y}\sum_{x\in X} g(x,y)\cdot f(x,y)$$

$$E[g(x,y)] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y)\cdot f(x,y)$$

مثال: توپ‌های آبی و قرمز و سبز (۳ توپ آبی،
۲ قرمز، ۳ سبز) آزمایش تصادفی
به صورت تصادفی خارج می‌کنیم. $E[XY]=?$

$f(x,y)$	$x=0$	$x=1$	$x=2$
$y=0$	$\frac{3}{28}$	$\frac{9}{28}$	$\frac{3}{28}$
$y=1$	$\frac{3}{14}$	$\frac{3}{14}$	0
$y=2$	$\frac{1}{28}$	0	0

X: تعداد توپ‌های آبی بیرون آمده از سبد
Y: 〃 〃 قرمز 〃 〃 〃

$$\sum_{x=0}^{2}\sum_{y=0}^{2} xy f(x,y) = 0\times 0\times f(0,0) + 0\times 1\times f(0,1) + 0\times 2\times f(0,2) + \dots + 2\times 2\times f(2,2) = \frac{3}{14}$$

مثال: برای متغیرهای تصادفی X و Y با تابع توزیع احتمال توأم زیر $E[\frac{Y}{X}]$ را به دست آورید.

$$f(x,y) = \begin{cases} \dfrac{x(1+3y^2)}{4} & 0<x<2 \quad 0<y<1 \\ 0 & \text{o.w} \end{cases}$$

Subject:

$$E\left[\frac{Y}{X}\right] = \int_0^1\int_0^2 \frac{y}{x}\cdot\frac{x(1+3y^2)}{4}\,dx\,dy$$

$$= \int_0^1 \frac{y+3y^3}{2}\,dy = \frac{5}{8}$$

$$\underset{\text{تابعی از x و y است}}{E[X]} = \int_{-\infty}^{+\infty} x.f(x)\,dx = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x\,f(x,y)\,dx\,dy$$

$$X \to f(x)$$
$$X \to E[X]$$
$$X \to E[(X-E[X])^2]$$

واریانس $\frac{\sum(x-\bar{x})^2}{n}$

میانگین مربعات فاصله از میانگین

واریانس متغیر تصادفی X به صورت زیر تعریف می‌شود:

$$\delta_x^2 = var[X] = E[(X-E[X])^2] = E[\underbrace{(X-\underbrace{\mu_x}_{E[x]})^2}_{g(x)}]$$

$$var[X] = \sum_{x\in X}\underbrace{(x-\mu_x)^2}_{g(x)} f(x)$$

$$var[X] = \int_{-\infty}^{+\infty}(x-\mu_x)^2 f(x)\,dx$$

$$var[X] = E[(X-\mu_x)^2] = E[X^2] - (E[X])^2$$

$$\int_{-\infty}^{+\infty} x^2 f(x)\,dx \qquad \left(\int_{-\infty}^{+\infty} x.f(x)\,dx\right)^2$$

Subject:

اثبات:

$$\sum_{x\in X}(x-\mu_x)^2 f(x) = \sum (x^2+\mu_x^2-2\mu_x x) f(x)$$

$$\sum_{x\in X} x^2 f(x) + \mu_x^2 \sum_{x\in X} f(x) - 2\mu \sum_{x\in X} x f(x)$$

$$E[X^2] + \mu_x^2 - 2\mu_x^2 = E[X^2] - \mu_x^2 = E[X^2] - (E[X])^2$$

جلسه یازدهم ۹۷/۱/۲۸

واریانس:

$$V[x] = E[(X-\mu_x)^2] = E[x^2] - \underbrace{E[x]^2}_{\mu_x^2}$$

مثال: اگر متغیر تصادفی X را تعداد قطعات خراب در یک نمونه تصادفی ۳ تایی از قطعات تولیدی یک کارخانه در نظر بگیریم که تابع توزیع جرم احتمال X به صورت زیر باشد، واریانس را بیابید.

x	۰	۱	۲	۳
$f(x)$	۰/۵۱	۰/۳۸	۰/۱	۰/۰۱

$$\sigma_x^2 = Var[x] = E[x^2] - E[x]^2 = (0/87) - (0/41)^2 = 0/49$$

$$E[x] = 0\times 0/51 + 1\times 0/38 + 2\times 0/1 + 3\times 0/01 = 0/41$$

$$E[x^2] = 0^2\times 0/51 + 1^2\times 0/38 + (2)^2\times 0/1 + (3)^2\times 0/01 = 0/87$$

$$\sum_x x^2 f(x)$$

* واریانس همواره مقداری مثبت است. $Var[x] \geq 0$

$Var[x] = 0$ به معنای این است که هیچ پراکندگی در داده ها نداریم.

Subject:

مثال: مقدار موجودی آب آشامیدنی برای یک شهر یک نفر برحسب مترمکعب یک متغیر تصادفی x با توزیع زیر است، میزان واریانس مقدار موجودی آب را بیابید.

$$f(x) = \begin{cases} 2(x-1) & 1 \le x < 2 \\ 0 & o.w \end{cases}$$

$$var[x] = E[x^2] - E[x]^2 = \frac{17}{6} - \left(\frac{5}{3}\right)^2 = \frac{1}{18}$$

$$E[x] = \int_1^2 x * 2(x-1)\, dx = \frac{5}{3}$$

$$E[x^2] = \int_1^2 x^2 * 2(x-1)\, dx = \frac{17}{6}$$

محدوده انتگرال همان محدوده x است.

نامساوی چبیشف (chebyshev)

$$\sigma_x = \sqrt{var[x]}$$

$$P(\mu_x - k\sigma_x < X < \mu_x + k\sigma_x) \ge 1 - \frac{1}{k^2} \quad *$$

نامساوی مارکوف

$$P(X \ge a) \le \frac{E[X]}{a} \quad *$$

* برای اثبات نامساوی چبیشف، از نامساوی مارکوف استفاده می‌کنیم

مثال: $P(X \ge 17) \le \frac{14}{17}$

$E[X] = 14$ با فرض

Subject:

کوواریانس

$$cov[X,Y] = E[(X-\mu_x).(Y-\mu_y)]$$

$cov(x,y) > 0$ ← x و y دارای همبستگی مثبت خطی هستند یعنی افزایش یکی با افزایش دیگری همراه است. (به طور کلی)

$cov(x,y) < 0$ ← X و Y دارای همبستگی خطی منفی هستند یعنی به طور کلی افزایش یک متغیر با کاهش دیگری همراه است.

$$cov[x,y] = E[xy - x\mu_y - y\mu_x + \mu_x\mu_y] =$$

$$E[XY] - \mu_y \underbrace{E[X]}_{\mu_x} - \mu_x \underbrace{E[Y]}_{\mu_y} + \mu_x\mu_y = E[XY] - \mu_x\mu_y =$$

$$\underbrace{E[XY]}_{} - E[X]E[Y]$$

$$\neq E[X].E[Y]$$

ضریب همبستگی $\rho(x,y) = \dfrac{cov(X,Y)}{\sigma_X \sigma_Y}$

همواره در بازه $[-1,1]$ است.

* اگر $cov(x,y) = 0$ باشد، نتیجه‌ای نمی‌توان گرفت. نمی‌توان گفت x و y مستقل اند.

ضریب همبستگی مقادیر بین ۱- تا ۱ دارد

$|\rho(x,y)| > 0{,}7$ ← همبستگی خطی قوی وجود دارد

$0{,}5 < |\rho(x,y)| < 0{,}7$ ← " " متوسط " "

$0{,}3 < |\rho(x,y)| < 0{,}5$ ← " " ضعیف " "

$0 < |\rho(x,y)| < 0{,}3$ ← " " تقریباً وجود ندارد

IDEA

Subject:

مثال: برای دو متغیر تصادفی با تابع توزیع احتمال توأم زیر cov بین این دو متغیر را بیابید.

$$f(x,y) = \begin{cases} \dfrac{x(1+3y^2)}{4} & 0<x<2 \quad 0<y<1 \\ 0 & o.w \end{cases}$$

$$cov[x,y] = E[XY] - E[X]E[Y] = \frac{5}{6} - \frac{5}{8} \times \frac{4}{3} = 0$$

$$E[XY] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \frac{xy \cdot x(1+3y^2)}{4}\,dx\,dy = \frac{5}{6}$$

$$E[X] = \int_{-\infty}^{+\infty} x f(x)\,dx = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x f(x,y)\,dx\,dy =$$

$$\int_0^1\int_0^2 x \cdot \frac{x(1+3y^2)}{4}\,dx\,dy = \frac{4}{3}$$

$$f(x) = \int_{-\infty}^{+\infty} f(x,y)\,dy$$

$$E[Y] = \int_0^1\int_0^2 \frac{y \cdot x(1+3y^2)}{4}\,dx\,dy = \frac{5}{8}$$

این دو متغیر به یکدیگر وابستگی خطی ندارند.

Subject:

۱۳۹۷/۲/۱۲

جلسه دوازدهم

$$cov(x,y) = E[(x-\mu_x)(y-\mu_y)] = E[XY] - E[X].E[Y]$$

اگر $cov(x,y) = 0 \Longrightarrow E[XY] = E[X].E[Y]$

لزوماً به معنای اینکه X و Y از هم مستقل هستند نیست. معادل این نیست

دقت شود: تعریف استقلال $\forall x,y \quad f(x,y) = f(x) f(y)$ اما اگر

X و Y از هم مستقل باشند آنگاه $cov(x,y)=0$

ـ ثابت کنید اگر X و Y از هم مستقل باشند آنگاه $cov(x,y)=0$ ؟

$$cov(x,y) = E[XY] - E[X]E[Y] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xy\, f(x,y)\,dx\,dy$$

$$-\left(\int_{-\infty}^{+\infty} x f_X(x)\,dx\right)\left(\int_{-\infty}^{+\infty} y f_Y(y)\,dy\right) = 0$$

چون از هم مستقل اند $f(x,y) = f(x) f(y)$

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xy\, f_X(x).f_Y(y) - \left(\int_{-\infty}^{+\infty} x f_X(x)\right)\left(\int_{-\infty}^{+\infty} y f_Y(y)\right) = 0$$

$$\int_{-\infty}^{+\infty} x f_X(x) \int_{-\infty}^{+\infty} y f_Y(y)$$

Subject:

مثال: در مثال توپ $cov(x,y)$ را بیابید.

$y \backslash X$	۰	۱	۲	$h(y)$
۰	$\frac{3}{28}$	$\frac{9}{28}$	$\frac{3}{28}$	$\frac{15}{28}$
۱	$\frac{3}{14}$	$\frac{3}{14}$	۰	$\frac{3}{7}$
۲	$\frac{1}{28}$	۰	۰	$\frac{1}{28}$
$g(x)$	$\frac{5}{14}$	$\frac{15}{28}$	$\frac{3}{28}$	

X: تعداد توپ آبی

Y: تعداد توپ قرمز

۳ آبی
۳ سبز
۲ قرمز

۲ توپ به تصادف انتخاب می شود

$$E[X] = \sum_{x=0}^{2} x.\overbrace{g(x)}^{f_X(x)} = 0 \times \frac{5}{14} + 1 \times \frac{15}{28} + 2 \times \frac{3}{28} = \frac{3}{4}$$

$$E[Y] = \sum_{y=0}^{2} y.h(y) = 0 \times \frac{15}{28} + 1 \times \frac{3}{7} + 2 \times \frac{1}{28} = \frac{1}{2}$$

$$E[XY] = \sum_{x=0}^{2} \sum_{y=0}^{2} xy.f(x,y) = 0 \times 0 \times \frac{3}{28} + 0 \times 1 \times \frac{9}{28} + \dots + 2 \times 2 \times 0 = \frac{3}{14}$$

$$cov(x,y) = E[XY] - E[X]E[Y] = \frac{-9}{56}$$

مقدار منفی چون افزایش تعداد توپ آبی با کاهش توپ قرمز همراه است و بالعکس!

$$\rho_{xy} = \frac{cov(x,y)}{\sigma_x \sigma_y} = \frac{-\frac{9}{56}}{\sqrt{\frac{45}{112}} \times \sqrt{\frac{9}{28}}} = \frac{-1}{\sqrt{5}} = -0.45$$

همبستگی خطی ضعیف

IDEA

Subject:

مثال: برای دو متغیر تصادفی X و Y تابع توزیع احتمال توأم به صورت زیر است، $cov(x,y) = ?$:

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

$$E[XY] = \int_0^1 \int_0^y xy \cdot \frac{1}{y}\, dx\, dy = \int_0^1 \frac{x^2}{2}\Big|_0^y dy = \int_0^1 \frac{y^2}{2} dy =$$

$$\frac{1}{6} y^3 \Big|_0^1 = \frac{1}{6}$$

$$E[X] = \int_0^1 \int_0^y x \cdot \frac{1}{y}\, dx\, dy = \int_0^1 \frac{1}{y} \left(\int_0^y x\, dx\right) dy$$

$$= \int_0^1 \frac{1}{2} y\, dy = \frac{1}{4}$$

$$E[Y] = \int_0^1 \int_0^y y \cdot \frac{1}{y}\, dx\, dy = \int_0^1 y\, dy = \frac{1}{2} y^2 \Big|_0^1 = \frac{1}{2}$$

$$cov(x,y) = \frac{1}{6} - \frac{1}{2} \times \frac{1}{4} = \frac{1}{24}$$

محاسبه امید ریاضی و واریانس تابعی از یک متغیر تصادفی:

$$Z = g(X) = aX + b \qquad E[Z] = E[aX+b] = aE[X] + b$$

$$E[Z] = \Sigma (ax+b) f_X(x) = a \sum_x x \cdot f_X(x) + \sum_x b f(x) \rightarrow b \sum_x f(x) = b$$

$$E[g(x,y)] = \sum_x \sum_y g(x,y) f(x,y)$$

$$E[g(x,y) \pm h(x,y)] = E[g(x,y)] \pm E[h(x,y)]$$

Subject:

$$E[g(x) \pm h(y)] = E[g(x)] \pm E[h(y)]$$

مثال: $E[X^2 + Y^3] = E[X^2] + E[Y^3]$

$$E[X - Y] = E[X] - E[Y]$$ ← هیچ شرایطی نیاز ندارد

$$\mathrm{var}[aX+b] = E[((aX+\cancel{b}) - \underbrace{E[aX+b]}_{aE[X]+\cancel{b}})^2] =$$

$$E[a^2(X-E[X])^2] = a^2 \underbrace{E[(X-\mu_x)^2]}_{\mathrm{var}[X]} = a^2\,\mathrm{var}[X]$$

$$\mathrm{var}[ax+by+c] = E[((ax+by+\cancel{c}) - \underbrace{E[ax+by+c]}_{-aE[X]-bE[Y]-\cancel{c}})^2]$$

$$E[(ax+by-a\mu_x - b\mu_y)^2] =$$

$$E[(ax-a\mu_x)^2 + (by-b\mu_y)^2 + \underbrace{2ab(ax-a\mu_x)(by-b\mu_y)}_{2ab[(x-\mu_x)(y-\mu_y)]}] =$$

$$a^2\,\mathrm{var}[X] + b^2\,\mathrm{var}[Y] + 2ab\,\mathrm{cov}(X,Y)$$

اگر X, Y از هم مستقل باشند → $\mathrm{var}[aX+by+c] = a^2\,\mathrm{var}[X] + b^2\,\mathrm{var}[Y]$

$\mathrm{cov}(x,y) = 0$

اگر برای یک متغیر تصادفی X مقدار $E[X] = \mu_x$ و $\mathrm{var}[X] = \sigma_x^2$ را در اختیار داشته باشیم

ولی تابع توزیع احتمال آن را در اختیار داشته باشیم آنگاه با استفاده از سبک تیلور می‌توان امید ریاضی و واریانس

Subject:

هر تابع دلخواه از X را بدست آورید:

$E[X] = \mu_x \qquad E[g(x)] = ?$

$var[X] = \sigma_x^2 \qquad var[g(x)] = ?$

بسط تیلور g(x) را حول نقطه $X = \mu_x$ می‌نویسیم:

$$g(X) = \sum_{n=0}^{\infty} \frac{(X-\mu_x)^n}{n!} \frac{\partial^n g(x)}{\partial x^n}\bigg|_{x=\mu_x} =$$

$$g(\mu_x) + (x-\mu_x)\, g'(x)\Big|_{x=\mu_x} + \frac{(x-\mu_x)^2}{2} g''(x)\Big|_{x=\mu_x}$$

از دو طرف رابطه E می‌گیریم $\longrightarrow$

$$E[g(x)] = E\left[g(\mu_x) + (x-\mu_x) g'(\mu_x) + \frac{(x-\mu_x)^2}{2} g''(\mu_x)\right]$$

$$= g(\mu_x) + \underbrace{g'(\mu_x) E[X-\mu_x]}_{0} + g''\left(\frac{\mu_x}{2}\right) \underbrace{E[(X-\mu_x)^2]}_{var[X]}$$

$$= g(\mu_x) + \frac{g''(\mu_x)}{2} var[X]$$

۹۷,۲,۴   جلسه سیزدهم

امید ریاضی و واریانس هر تابع دلخواه از یک متغیر تصادفی

$$\begin{cases} var[g(x)] = g'^2(\mu_x)\, var[X] \\ E[g(x)] = g(\mu_x) + \dfrac{var[X]}{2} g''(\mu_x) \end{cases}$$

Subject:

مثال :
$E[X] = \mu_x = 15$ $\quad g(x) = \sqrt{20x}$

$var[X] = 9$

$$E(\sqrt{20x}) = \sqrt{20 \times 15} + \frac{9}{2} \times \frac{-\sqrt{20}}{4} x^{-\frac{3}{2}} \Big|_{x=15}$$

$$var[\sqrt{20x}] = \left(\frac{\sqrt{20}}{2} (15)^{-\frac{1}{2}} \right)^2 = 9$$

مثال: اگر برای متغیر X مقدار $E[X]$ برابر μ_x و $Var[X]$ برابر σ_x^2 باشد

$E[e^x]$ و $var[e^x]$ را تخمین بزنید.

$$E[e^x] = e^{\mu_x} + \frac{\sigma_x^2}{2} \cdot e^{-\mu_x}$$

$$var[e^x] = e^{2\mu_x} \cdot \sigma_x^2$$

"پایان ترم"

فصل ۵

توزیع های گسسته معروف

$f(x)$ _ $E[X]$ _ $var[X]$

۱_ برنولی

۲_ چند جمله ای

۳_ فوق هندسی

۴_ پواسون

Subject:

متغیر تصادفی برنولی (دارای توزیع جرم احتمال برنولی)

آزمایش برنولی : یک آزمایش تصادفی است که دو خروجی آن اهمیت دارد (شکست و موفقیت)

(P : احتمال موفقیت و $q = 1-P$: احتمال شکست)

متغیر تصادفی برنولی : تعداد موفقیت در یک بار آزمایش برنولی (X) (نگاشت خروجی به اعداد حقیقی)

تابع توزیع متغیر تصادفی برنولی

$$f(x) = f(x;P) = \begin{cases} P & x=1 \\ 1-P=q & x=0 \\ 0 & \text{o.w} \end{cases}$$

پارامتر توزیع برنولی

$$E[X] = 1\times P + 0\times(1-P) = P$$

امید برنولی

$$\mathrm{var}[X] = \underbrace{E[X^2]}_{\underbrace{1^2 P + 0^2(1-P)}_{P}} - E[X]^2 = P - P^2 = P(1-P) = qP$$

متغیر تصادفی دوجمله‌ای (Binomial) :

آزمایش دوجمله‌ای : نتیجه n بار تکرار آزمایش برنولی .

نتیجه هر بار آزمایش برنولی از دفعات دیگر مستقل است .

Subject:

متغیر تصادفی دو جمله‌ای: تعداد موفقیت‌ها در n بار تکرار آزمایش

تابع توزیع متغیر تصادفی دو جمله‌ای

$$\underbrace{b\,(x\,;\,n,p)}_{\text{پارامترهای توزیع دوجمله‌ای}} = \begin{cases} \binom{n}{x} p^{x} (1-p)^{n-x} = \binom{n}{x} p^{x} q^{n-x} & x=0,1,\dots,n \\ 0 & \text{o.w} \end{cases}$$

$(p+q)^n$: دلیل ۲ جمله‌ای بودن

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

در دوجمله‌ای

$$E[X] = [Y_1 + Y_2 + \dots + Y_n] = E[Y_1] + E[Y_2] + \dots + E[Y_n] = nq$$

حاصل هر بار آزمایش pq pq

$$var[X] = var[Y_1 + \dots + Y_n] = var[Y_1] + \dots + var[Y_2] = npq$$

برنولی — از هم مستقل است

مثال: احتمال زنده ماندن افرادی که به بیماری سرطان مبتلا می‌شوند ۰٫۴ است از بین ۱۵ نفر که به سرطان مبتلا شده‌اند چقدر احتمال دارد بیش از ۱۰ نفر زنده بمانند؟

$$P(X \geq 10) = \sum_{10}^{15} \binom{15}{x} (0.4)^{x} (0.6)^{15-x} = 1 - CDF(9) = 0.034$$

تابع توزیع تجمعی

$$CDF(x) = P(X \leq x)$$

توزیع ۲ جمله‌ای $b(15, 0.4)$

Subject:

$$P(۳ \le x \le ۸) = \sum_{۴}^{۸} \binom{۱۵}{x} (۰/۴)^{x} (۰/۶)^{۱۵-x} = CDF(۸) - CDF(۴)$$

$$= ۰/۹۰۵ - ۰/۰۲۷۱ = ۰/۸۷$$

متغیر تصادفی پواسون :

متغیر تصادفی پواسون درباره احتمال رخداد یک حادثه یا رویداد به تعداد مشخصی در $\underbrace{\text{بازه زمانی}}_{\text{واحد زمانی}}$ یا $\underbrace{\text{فاصله مکانی}}_{\text{واحد مکان}}$ مشخص رخ دهد صحبت می‌کند.

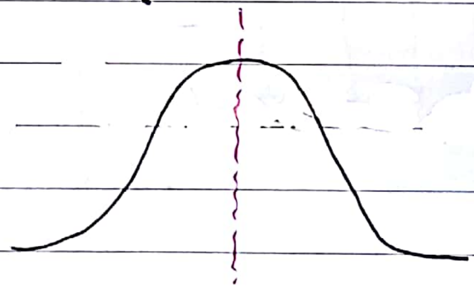
تعداد رخدادها در هر بازه زمانی (مکانی) از بازه دیگر مستقل است.

تعداد رخدادها در بازه‌های زمانی متناسب با طول بازه است.

چند برابر واحد زمان

$$P(X ; \lambda t) = \frac{e^{-\lambda t} (\lambda t)^{x}}{x!}$$

میانگین تعداد دفعات رخداد در واحد زمان



ده دقیقه به طور میانگین ۷ نفر از درب ولیعصر رد می‌شوند در ۲۰ دقیقه در ۳۰ دقیقه ...

$$P(۳۰ ; ۲۰۰) = \frac{e^{-۲۰} (۲۰۰)^{-۳۰}}{۳۰!}$$

Subject:
۹۷/۲/۹

جلسه چهاردهم

شبکه های کامپیوتری ←

توزیع پواسون: (Poisson) ← نظریه صف

توزیع پواسون درباره احتمال تعداد دفعات رخداد یک واقعه در یک واحد زمان یا واحد مکان با داشتن میانگین تعداد دفعات رخداد آن واقعه در واحد زمان صحبت می کند.

پارامتر توزیع (با داشتن این پارامتر می توان احتمال حالات مختلف را حساب کرد)

$$P(x; \lambda t) = \frac{(\lambda t)^x e^{-\lambda t}}{x!}$$

← احتمال اینکه x دفعه واقعه در بازه λt رخ دهد.

x: تعداد دفعات رخداد در واحد زمان (دلخواه)

λ: میانگین دفعات رخداد در واحد زمان

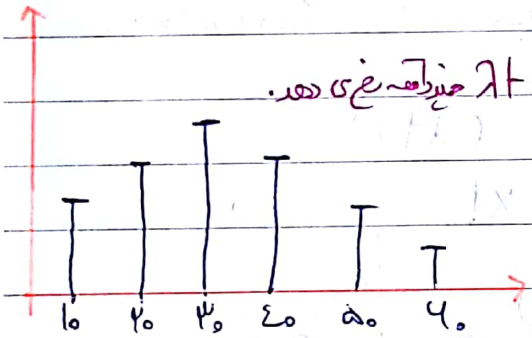
t: چند برابر واحد زمان

به طور میانگین در t برابر واحد زمان λt رخداد اتفاق می افتد.

$E[X] = \lambda t$

$Var[X] = \lambda t$

به طور میانگین در بازه λt چند واقعه رخ می دهد.



مثال: تعداد خطاهای یک تایپیست در هر صفحه به طور میانگین برابر ۴ است.

الف) با چه احتمالی این تایپیست ۶ خطا در صفحه خواهد داشت؟

ب) با چه احتمالی در ۱۰۰۰ صفحه ای که این تایپیست تایپ می کند ۱۰۰ خطا خواهد داشت؟

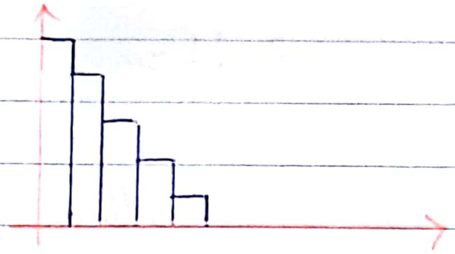
الف) $P(6; 4 \times \underbrace{1}_{\text{یک صفحه}}) = \frac{(4)^6 \times e^{-4}}{6!}$

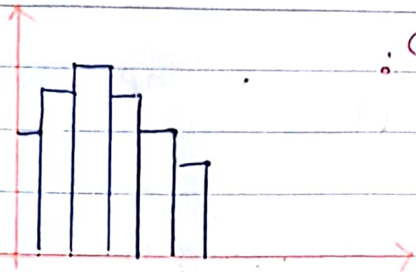
در واحد یک صفحه

ب) $P(100; \underbrace{4}_{\lambda} \times \underbrace{1000}_{\text{۱۰۰۰ صفحه یا ۱۰۰۰ برابر واحد زمان}}) = \frac{(4000)^{100} \times e^{-4000}}{100!}$

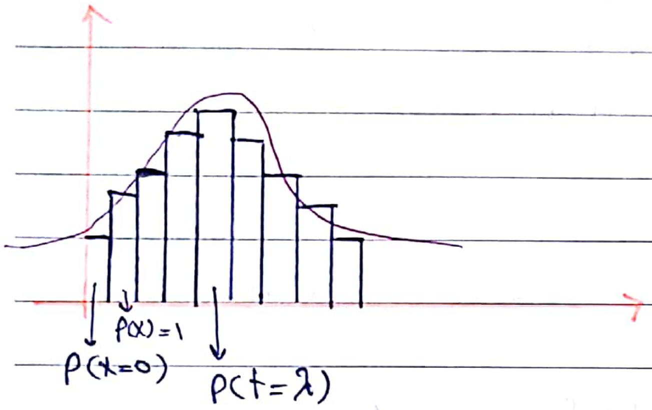
Subject:

شکل پواسون :





اگر λ به اندازه کافی بزرگ باشد : شکل منظم‌تر و منسجم‌تری می‌گیرد



تقریب مقادیر احتمالات یک توزیع براساس توزیع دیگر :

① تقریب توزیع دوجمله‌ای با توزیع پواسون

مثال: احتمال خوب شدن یک بیماری برابر p ، n نفر مبتلا هستند چقدر احتمال دارد x نفر خوب شوند؟

$$b(x;n,p) = \binom{n}{x} p^x (1-p)^{n-x} \longrightarrow E[X] = np$$

توزیع دوجمله‌ای

$$var[X] = npq \quad (q = 1-p)$$

$$E[X] = E[Y] \qquad var[X] = var[Y]$$

دوجمله‌ای — توزیع پواسون به کار گرفته شده در تقریب

Subject:

$$\begin{cases} np = \lambda t \\ np\underbrace{q}_{1-p} = \lambda t \end{cases} \Rightarrow np = np(1-p) \longrightarrow np \text{ !! } $$ تقریبا

برای این منظور نیاز داریم پارامترهای توزیع پواسون را به گونه‌ای تعیین کنیم که توزیع پواسون تابع توزیع احتمالی شبیه به توزیع دوجمله‌ای داشته باشد.

شرایط تقریب زدن ←

$n \to \infty$

$p \to 0 , 1-p \to 1$

np باید به سمت یک عدد ثابت میل کند

$\boxed{\lambda t = np}$ انتخاب

مثال: احتمال خرابی IC هایی که یک کارخانه تولید می‌کند ۰٫۰۰۱ است. اگر ۸۰۰۰ تا از این IC ها تولید شود با چه احتمالی کمتر از ۷ تا از این IC ها خراب است؟

$$\sum_{x=0}^{6} b(x; 8000, 0.001) = \sum_{x=0}^{6} \binom{8000}{x} (0.001)^x (0.999)^{8000-x}$$

⇓ تقریب با پواسون

$$\lambda t = np = 8000 \times \frac{1}{1000} = 8 \longrightarrow \sum_{x=0}^{6} \frac{e^{-8} 8^x}{x!} = CDF(6)$$

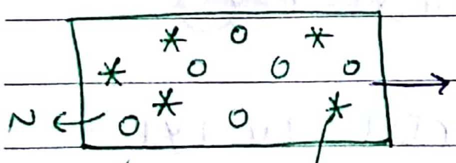
پارامتر توزیع پواسون برابر np توزیع دوجمله‌ای قرار داده می‌شود

توزیع پواسون با $\lambda t = 8$

توزیع فوق هندسی: (اَبَر)

x: تعداد اشیایی که ویژگی * را دارند.

آزمایش تصادفی



n آیتم به تصادف بدون جایگذاری خارج می‌کنیم

k تا از N دارای ویژگی (*) هستند

توزیع دوجمله‌ای این اتفاق با جایگذاری رخ می‌دهد.

آزمایش توزیع فوق هندسی (داتی) :

آزمایش تصادفی‌ای را در نظر بگیرید که از میان N شیء که k تای آن دارای ویژگی ⊛ ، n شیء به تصادف و بدون جایگذاری

انتخاب می‌کنیم . متغیر تصادفی X را تعداد اشیای از نمونه n تایی در نظر می‌گیریم که دارای ویژگی ⊛ هستند ، آنگاه X دارای

توزیع فوق هندسی خواهد بود .

$$f(x) = h(x; N, k, n) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}$$

$$E[X] = \frac{k}{N} \cdot n \quad , \quad var[X] = \frac{N-n}{N-1} \cdot n \cdot \frac{k}{N} \cdot \left(1 - \frac{k}{N}\right)$$

اگر با جایگذاری بود معادل یک توزیع دوجمله‌ای بود :

$$b(x; n, p) \quad n = n \,,\quad p \to \frac{k}{N}$$

جلسه یازدهم ۹۷/۲/۱۴

⁎ اگر $\frac{N-n}{N-1} \to 1$ میل کند ، آنگاه $\underset{\text{توزیع دوجمله‌ای}}{var[X]} = \underset{\text{فوق هندسی}}{var[Y]}$ → شرایط امکان تقریب فوق هندسی با دوجمله‌ای

⇐ اگر $N \to \infty$ و $\frac{n}{N} \to 0$ می‌توان توزیع فوق هندسی $h(x; N, n, k)$ را می‌توان با یک

توزیع دوجمله‌ای $b(x; n, \frac{k}{n})$ تقریب زد ($p = \frac{k}{n}$)

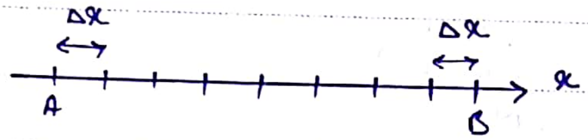
میانگین پارامتر توزیع دوجمله‌ای معادل $n \cdot \frac{k}{n} = n \cdot p$

فصل ۶

توزیع های پیوسته معروف (یکنواخت، نرمال، گاما، نمایی)

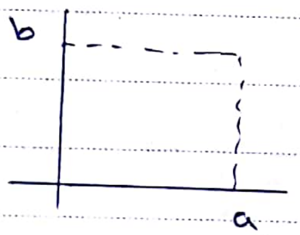
توزیع یکنواخت:

متغیر تصادفی X دارای توزیع یکنواخت در بازه $[A,B]$ است، اگر تابع توزیع آن به صورت زیر باشد:



$$f(x) = \begin{cases} \frac{1}{B-A} & A \leq x \leq B \\ 0 & o.w \end{cases}$$

$E[X] = \frac{A+B}{2}$ وسط بازه $\qquad var[X] = \frac{(B-A)^2}{12}$



$f(x,y) = \frac{1}{\text{مساحت}} \qquad x,y \in R$

نرمال (گوسی):

چگالی احتمال

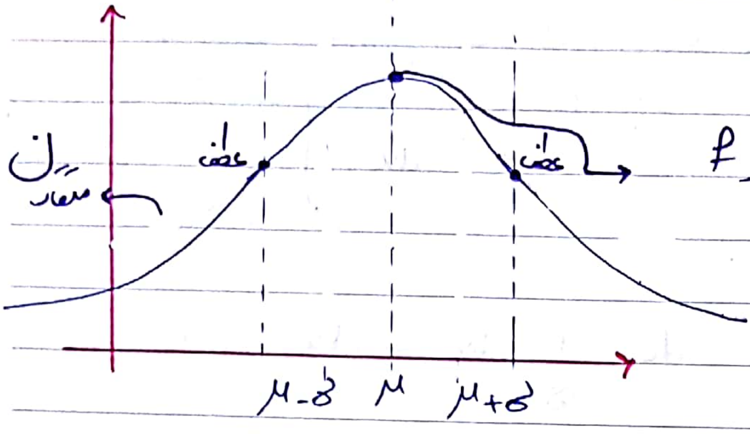
$N(x;\mu,\sigma)$

متغیر تصادفی X با تابع توزیع زیر متغیر تصادفی نرمال نامیده می شود.

$$N(x,\mu,\sigma) = f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$E[X] = \mu \qquad var[X] = \sigma^2$

Subject:



مد: تابع توزیع بیشترین مقدار را دارد.

محور میانگین با مد برابر است و بیشترین مقدار f را دارد.

وقتی σ ۳ دور می‌شیم احتمال صفر می‌شود.

مثال: اگر بیست آزمون میانگین نمرات ۵۰ و انحراف معیار آن‌ها ۱۰ باشد و بدانیم نمرات توزیع نرمال دارند چقدر احتمال دارد فردی نمره‌ای بالای ۷۰ کسب کند؟

$$P(70 \leq x) = \int_{70}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = 1 - CDF(70) \xrightarrow{N(50,10)} =$$

$$1 - CDF\left(\frac{70-50}{10}\right) = 1 - \underbrace{CDF(2)}_{\text{استاندارد } N(0,1)} \xrightarrow{\text{از روی جدول}} 1 - 0{,}9772 = 0{,}0228$$

توزیع نرمال استاندارد:

متغیر تصادفی x با تابع توزیع نرمال که $\mu=0$ و $\sigma=1$ باشد دارای توزیع نرمال استاندارد است

$$N(x;0,1) = \underbrace{N(0,1)}_{\text{نرمال استاندارد}}$$

Subject:

*نکته: بامحاسبات هر توزیع نرمال دلخواه، می‌توان به محاسبات توزیع نرمال استاندارد تبدیل کرد.

$$P(\alpha < X < \beta) = \int_{\alpha}^{\beta} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = \int_{z_1 = \frac{\alpha-\mu}{\sigma}}^{z_2 = \frac{\beta-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz$$

$N(\mu,\sigma)$ توزیع نرمال

$$z = \frac{x-\mu}{\sigma} \qquad dz = \frac{1}{\sigma}dx \Rightarrow dx = \sigma dz$$

توزیع نرمال استاندارد

$$P(\alpha < x < \beta) = CDF\left(\frac{\beta-\mu}{\sigma}\right) - CDF\left(\frac{\alpha-\mu}{\sigma}\right)$$

$N(\mu,\sigma)$, $N(0,1)$, $N(0,1)$

$$P(\alpha < x) = 1 - CDF\left(\frac{\alpha-\mu}{\sigma}\right) \qquad P(x < \beta) = CDF\left(\frac{\beta-\mu}{\sigma}\right)$$

جلسه شانزدهم ۹۷,۲,۱۸

تقریب توزیع دوجمله‌ای با استفاده از توزیع نرمال:

$$b(x; n, p) \begin{cases} E[X] = np \\ var[X] = np(1-p) \end{cases}$$

$$N(y; \mu, \sigma) \begin{cases} \mu = E[y] = E[X] = np \\ \sigma^2 = var[y] = var[X] = np(1-p) \end{cases}$$

Subject:

توزیع دوجمله‌ای را در شرایطی که $n \to \infty$ برای دقت خوب p یا q نزدیک ۰ یا ۱ نباشند.

p و $\frac{q}{1-p}$ مقداری نزدیک $\frac{1}{2}$ داشته باشند و n نسبتاً بزرگ باشد.

می‌توان توسط یک توزیع نرمال با میانگین $\mu = np$ و انحراف معیار $\sigma = \sqrt{npq} = \sqrt{np(1-p)}$

تقریب زد.

مثال: اگر یک بیماری احتمال خوب شدن ۰٫۴ باشد و ۱۵ نفر به این بیماری مبتلا باشند چقدر

احتمال دارد حداکثر ۶ نفر از آنها بهبود یابند؟

$$\sum_{x=0}^{6} b(x; 15, 0{,}4) = \sum_{x=0}^{6} \binom{15}{x} (0{,}4)^x (0{,}6)^{15-x}$$

نیم واحد بیشتر

$$= \int_{-\infty}^{6{,}5} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = CDF\left(\frac{6{,}5-6}{\sqrt{3{,}6}}\right)$$

$\mu \to np = 6$

$\sigma \to \sqrt{npq} = \sqrt{3{,}6}$

ب) چقدر احتمال دارد دقیقاً ۷ نفر بهبود یابند؟

نیم واحد بیشتر

$$\int_{6{,}5}^{7{,}5} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-6}{\sqrt{3{,}6}}\right)^2} dx = CDF\left(\frac{7{,}5-6}{\sqrt{3{,}6}}\right) - CDF\left(\frac{6{,}5-6}{\sqrt{3{,}6}}\right)$$

نیم واحد کمتر

توزیع گاما:

متغیر تصادفی X با تابع توزیع زیر را متغیر تصادفی گاما می‌نامند.

Subject:

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

پارامترهای توزیع
$\alpha>0 , \beta>0$

یادآوری: تابع گاما:
$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx \quad \alpha>0$$

$$\Gamma(1) = 1 \qquad \Gamma(\tfrac{1}{2}) = \sqrt{\pi} \qquad \Gamma(n) = (n-1)!$$

طبیعی

توزیع پواسون: به طور میانگین یک واقعه λ بار در واحد زمان رخ می دهد احتمال اینکه x واقعه در واحد زمان رخ دهد:

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

توزیع گاما:

درباره مدت زمانی که طول می کشد تا α وقوع از رخداد که تعداد رخداد آنها دارای توزیع پواسون است و به طور میانگین λ دفعه در واحد زمان رخ می دهد، صحبت می کند که $\beta = \frac{1}{\lambda}$ است.

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

میانگین مدت زمانی که طول می کشد تا اولین واقعه رخ دهد یا میانگین فاصله زمانی بین دو وقوع متوالی است.

فاصله زمانی بین دو وقوع متوالی از رخداد چقدر است؟
=
چقدر طول می کشد تا اولین واقعه رخ دهد

* وقتی $\alpha=1$ به توزیع گاما توزیع نمایی گویند.

IDEA

Subject:

مثال: اگر به یک مرکز تلفن به طور میانگین در هر ساعت ۱۰ تلفن ارجاع شود چقدر احتمال دارد که فاصله زمانی بین دو تلفنی که به این مرکز ارجاع می‌شود بین ۱۰ و ۲۰ دقیقه باشد؟

λ واحد زمان

$\beta = \frac{1}{10}$

$$\int_{\frac{10}{60}}^{\frac{20}{60}} 10 e^{-10x}\, dx$$

ب) چقدر احتمال دارد که مدت زمانی که طول می‌کشد تا پنجمین تلفن به مرکز ارجاع می‌شود، کمتر از ۴۰ دقیقه باشد؟

$$\int_0^{\frac{40}{60}=\frac{2}{3}} \frac{1}{\left(\frac{1}{10}\right)^5 \Gamma(5)} x^4 e^{-10x}\, dx$$

مثال: در برخی از موارد مدت زمانی که طول می‌کشد تا یک قطعه الکترونیکی عمل کند را با توزیع نمایی مدل می‌کنند.

$$f(t) = \frac{1}{\beta} e^{-\frac{t}{\beta}} \quad t>0 \longrightarrow P(X>t+S \mid X>t) = P(X>S)$$

طول عمر قطعه

اثبات:

$$P(X>t+S \mid X>t) = \frac{P(X>t+S \cap X>t)}{P(X>t)} = \frac{P(X>t+S)}{P(X>t)} = \frac{\int_{t+S}^{+\infty} \frac{1}{\beta} e^{-\frac{x}{\beta}}\, dx}{\int_{t}^{+\infty} \frac{1}{\beta} e^{-\frac{x}{\beta}}\, dx}$$

$$= \frac{e^{-\frac{t+S}{\beta}}}{e^{-\frac{t}{\beta}}} = e^{-\frac{S}{\beta}} = \int_S^{+\infty} \frac{1}{\beta} e^{-x/\beta}\, dx$$

خاصیت بدون حافظه بودن توزیع نمایی:

$$P(X>t+S \mid X>t) = P(X>S)$$

$$P(X>t+S) = P(X>S) \cdot P(X>t)$$

IDEA

Subject:
۹۷ | ۲ | ۲۲

جلسه هفدهم

فصل هفتم

چگونه تابع توزیع تابعی از یک متغیر تصادفی را به دست آوریم؟

مسئله را برای این حالات مختلف حل می‌کنیم:

① گسسته / پیوسته

② یک متغیر تصادفی / چند متغیر تصادفی

③ تابع تبدیل یک به یک / تابع تبدیل غیر یک به یک

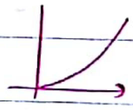
X	$f_X(x)$	X	$E[X]$	$var[X]$
$y=u(x)$	$f_Y(y)$	$y=g(x)$	"	"

تابع تبدیل

X	$X=1$	$X=2$	$X=3$	y	1	4	9	→ تعریف دامنه Y
$f_X(x)$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{2}$	$f_Y(y)$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{2}$	کنیم

$y=x^2$

با توجه به دامنه X یک نگاشت یک به یک از X به Y وجود دارد.



متغیر تصادفی گسسته (برای یک متغیر):

اگر برای متغیر تصادفی گسسته X تابع توزیع احتمال به صورت $f_X(x)$ باشد و متغیر تصادفی Y به صورت $y=u(x)$ تعریف شود که u تابعی یک به یک از X به Y باشد، آنگاه تابع توزیع جرم احتمال Y به صورت زیر می‌باشد.

$$f_Y(y)=f_X(u^{-1}(y))$$

X را برحسب Y حساب می‌کنیم

$$y=u(x) \Rightarrow x=u^{-1}(y)$$

Subject:

برای چند متغیره

اگر برای دو متغیر تصادفی پیوسته X_1 و X_2 تابع توزیع احتمال توأم X_1 و X_2 به صورت $f_{X_1,X_2}(x_1,x_2)$ باشد و متغیرهای تصادفی Y_1 و Y_2 به صورت $Y_1=u_1(X_1,X_2)$ و $Y_2=u_2(X_1,X_2)$ تعریف شوند به طوری که یک نگاشت یک به یک بین زوج مرتب‌های (x_1,x_2) و (y_1,y_2) وجود داشته باشد آنگاه تابع توزیع احتمال توأم Y_1 و Y_2 به صورت زیر خواهد بود:

$$f_{Y_1,Y_2}(y_1,y_2) = f_{X_1,X_2}\left(\underbrace{w_1(y_1,y_2)}_{\text{را برحسب } y_1, y_2 \text{ بدست می‌آوریم } x_1},\ \underbrace{w_2(y_1,y_2)}_{\text{را برحسب } y_1, y_2 \text{ بدست می‌آوریم } x_2}\right)$$

که در آن

$$\begin{cases} x_1 = w_1(y_1,y_2) \\ x_2 = w_2(y_1,y_2) \end{cases}$$

است.

مثال: X_1 و X_2 از هم مستقل اند و X_1: تعداد تلفن‌هایی که در واحد زمان زده می‌شود $\lambda_1=10$

X_2: تعداد فکس‌هایی که در واحد زمان زده می‌شود $\lambda_2=2$

X_1+X_2 نیز توزیع پواسن است و $\lambda_1=12$

* نکته: مجموع دو متغیر تصادفی مستقل با توزیع‌های پواسن با پارامترهای λ_1 و λ_2 یک متغیر با توزیع پواسن با پارامتر $\lambda_1+\lambda_2$ است.

Subject:

مثال: اگر X_1 و X_2 دو متغیر تصادفی مستقل با توزیع پواسون با پارامترهای λ_1 و λ_2 باشند تابع توزیع $Y = X_1 + X_2$ را بدست آورید.

① تابع توزیع احتمال توأم X_1 و X_2 را حساب می‌کنیم چون مستقل هستند

$$f_{X_1, X_2}(x_1, x_2) = \frac{e^{-\lambda_1} \lambda_1^{x_1}}{x_1!} \cdot \frac{e^{-\lambda_2} \lambda_2^{x_2}}{x_2!}$$

$$= \frac{e^{-(\lambda_1+\lambda_2)} \lambda_1^{x_1} \lambda_2^{x_2}}{x_1! \, x_2!} \qquad \begin{cases} x_1 = 0, 1, 2, \dots \\ x_2 = 0, 1, 2, \dots \end{cases}$$

② یک متغیر تصادفی دیگر برحسب X_1 و X_2 اضافه می‌کنیم $Y_2 = X_2$

③ بدست آوردن X_1 و X_2 برحسب Y_1 و Y_2

$$\begin{cases} X_1 = Y_1 - Y_2 \\ X_2 = Y_2 \end{cases}$$

④ جایگذاری در $f_{X_1, X_2}(x_1, x_2)$

$$f_{Y_1, Y_2}(y_1, y_2) = \frac{e^{-(\lambda_1+\lambda_2)} \lambda_1^{(y_1-y_2)} \cdot \lambda_2^{y_2}}{(y_1-y_2)! \, y_2!}$$

* دقت: در این مرحله هیچ x ای نباید باشد!

⑤ حتماً دامنه Y_1 و Y_2 را مشخص نمایید.

$$\text{معادل} * \begin{cases} y_1 = y_2, y_2+1, \dots \\ y_2 = 0, 1, 2, \dots \end{cases}$$

$$\begin{cases} y_1 = 0, 1, 2, \dots \\ y_2 = 0, 1, 2, \dots, y_1 \end{cases}$$

Subject:

۴) تابع توزیع حاشیه‌ای y_1 را حساب کنید.

$$f_{Y_1}(y_1) = \sum_{y_2=0}^{y_1} \frac{e^{-(\lambda_1+\lambda_2)} \lambda_1^{(y_1-y_2)} \lambda_2^{y_2}}{(y_1-y_2)!\, y_2!} \times \frac{y_1!}{y_1!}$$

$$= \frac{e^{-(\lambda_1+\lambda_2)}}{y_1!} \underbrace{\sum_{y_2=0}^{y_1} \underbrace{\frac{y_1!}{(y_1-y_2)!\, y_2!}}_{\binom{y_1}{y_2}} \lambda_1^{(y_1-y_2)} \lambda_2^{y_2}}_{(\lambda_1+\lambda_2)^{y_1}}$$

$$= \frac{e^{-(\lambda_1+\lambda_2)} (\lambda_1+\lambda_2)^{y_1}}{y_1!}$$ توزیع پواسون با پارامتر $\lambda_1+\lambda_2$

متغیر تصادفی پیوسته

x	1	2	3	y	1	4	9
$f_X(x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{3}$	$f_Y(y)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{3}$

$y = x^2$

یک به یک باشد.

$y = u(X)$

متغیر تصادفی پیوسته

اثبات در بیست.

$f_Y(y) = ?$ $\qquad CDF_Y(y) = ?$

اگر صعودی باشد

$F_Y(y) = P(Y \leq y) = P(u(X) \leq y)$

$P(X \leq u^{-1}(y)) = F_X(u^{-1}(y))$ $\qquad CDF_X$ است از نقطه $u^{-1}(y)$

$$\Rightarrow f_Y(y) = \frac{\partial u^{-1}(y)}{\partial y} \cdot f_X(u^{-1}(y))$$

+

Subject:

اگر نزولی باشد

$F_Y(y) = P(Y \leq y) = P(u(X) \leq y)$

$P(X \geq u^{-1}(y)) = 1 - F_X(u^{-1}(y))$ $\qquad \frac{\partial F(x)}{\partial x} \quad \frac{\partial x}{\partial y}$

$\Rightarrow f_Y(y) = -\frac{\partial u^{-1}(y)}{\partial y} \cdot f_X(u^{-1}(y))$

* دقت: فقط در حالت پیوسته می‌توان مشتق گرفت (به خاطر CDF)

اگر برای متغیر تصادفی پیوسته X تابع توزیع احتمال X برابر $f_X(x)$ باشد و متغیر تصادفی Y به صورت $Y = u(X)$ تعریف شود که u تابعی یک به یک باشد آنگاه تابع توزیع Y به صورت زیر خواهد بود:

$$f_Y(y) = \left|\frac{\partial u^{-1}(y)}{\partial y}\right| f_X(u^{-1}(y))$$

$u^{-1}(y)$ متعلق مشتق

مثال: متغیر تصادفی پیوسته X دارای تابع توزیع احتمال زیر است تابع توزیع احتمال $Y = 2X - 3$ را بیابید.

$$f_X(x) = \begin{cases} \frac{x}{12} & 1 < x < 5 \\ 0 & \text{o.w} \end{cases}$$

$X = \frac{y+3}{2}$

① X را برحسب Y حساب می‌کنیم

② مشتق X را نسبت به Y حساب می‌کنیم

③ دامنه Y را به دست می‌آوریم

$$f_Y(y) = \begin{cases} \dfrac{\frac{y+3}{2}}{12} \left|\frac{1}{2}\right| & \boxed{-1 < y < 7} \\ 0 & \text{o.w} \end{cases}$$

$\frac{\partial x}{\partial y} = \frac{1}{2}$

$\begin{cases} y = 2x - 3 \\ 1 < x < 5 \end{cases} \Rightarrow -1 < y < 7$

۹۷/۲/۲۵ — جلسه هجدهم

Random variable Transformation

بدست آوردن تابع توزیع احتمال تابعی از یک متغیر تصادفی

$X \longrightarrow f_X(x)$

تبدیل

$Y = u(X) \longrightarrow f_Y(y)$

تابع u یک به یک باشد

$X = u^{-1}(y)$

$\left| \frac{\partial x}{\partial y} \right|$

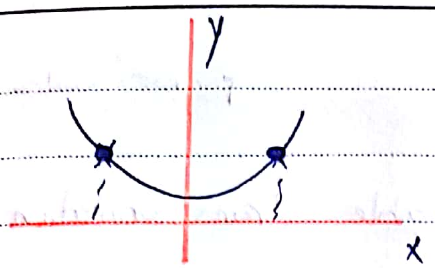
اگر تابع تبدیل یا تابع u غیر یک به یک باشد:

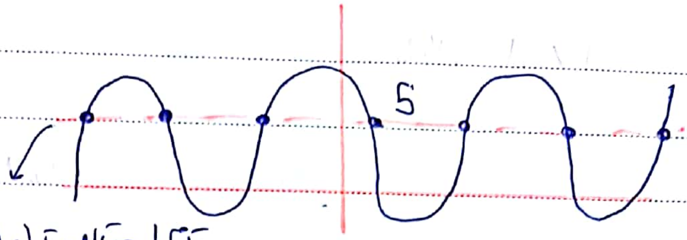
X گسسته باشد: $y = x^2 \leftarrow$ تابع در دامنه x یک به یک نیست

X	-2	-1	0	1	2	3
$f_X(x)$	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{2}{6}$	$\frac{1}{4}$

Y	0	1	4	9
$f_Y(y)$	$\frac{1}{12}$	$\frac{1}{12}+\frac{1}{12}$	$\frac{1}{6}+\frac{2}{6}$	$\frac{1}{4}$

برای y های ۴،۱
$$\begin{cases} X = \sqrt{y} \\ X = -\sqrt{y} \end{cases}$$





تمام احتمالات را با هم جمع می کنیم

احتمال $P(Y=y)$ برابر است با مجموع احتمالات نقاطی از X که $y=u(x_i)$ برابر y است.

$$P(Y=y) = \sum_{x_i \mid y = u(x_i)} P(X=x_i)$$

به طوری که

متغیر تصادفی پیوسته (برای چند متغیره یک به یک است)؟

اگر متغیرهای X_1 و X_2 متغیرهای تصادفی پیوسته با تابع توزیع احتمال توأم $f_{X_1,X_2}(x_1,x_2)$ باشند و متغیرهای تصادفی Y_1 و Y_2 به صورت $Y_1=u_1(X_1,X_2)$ و $Y_2=u_2(X_1,X_2)$ تعریف شوند به طوری که u_1 و u_2 یک نگاشت یک به یک بین زوج مرتب های (X_1,X_2) و (Y_1,Y_2) فراهم آورند آنگاه تابع توزیع احتمال توأم Y_1 و Y_2 به صورت زیر خواهد بود:

$$f_{Y_1,Y_2}(y_1,y_2) = f(w_1(y_1,y_2), w_2(y_1,y_2)) \cdot |J|$$

قدر مطلق دترمینان ماتریس ژاکوبین

که w_1 و w_2 توابعی هستند که x_1 و x_2 را از روی y_1 و y_2 محاسبه می‌کنند

$$\begin{cases} x_1 = w_1(y_1, y_2) \\ x_2 = w_2(y_1, y_2) \end{cases}$$

و ماتریس ژاکوبین (ژاکوبی) هم به ماتریس زیر گفته می‌شود.

$$J = \begin{bmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{bmatrix}$$

مثال: فرض کنید برای دو متغیر تصادفی x_1 و x_2 تابع توزیع احتمال توأم به صورت زیر است:

اگر $y_1 = x_1^2$ و $y_2 = x_1 \cdot x_2$ را در نظر بگیریم تابع توزیع احتمال توأم y_1 و y_2 را بدست آورید.

$$f(x_1, x_2) = \begin{cases} 4x_1 x_2 & 0 < x_1 < 1 \text{ و } 0 < x_2 < 1 \\ 0 & \text{O.W} \end{cases}$$

1) محاسبه x_1 و x_2 برحسب y_1 و y_2

$$\begin{cases} x_1 = \sqrt{y_1} \\ x_2 = \frac{y_2}{\sqrt{y_1}} \end{cases}$$

Subject:
Year. Month. Date. ( )

۲) محاسبه ماتریس ژاکوبی

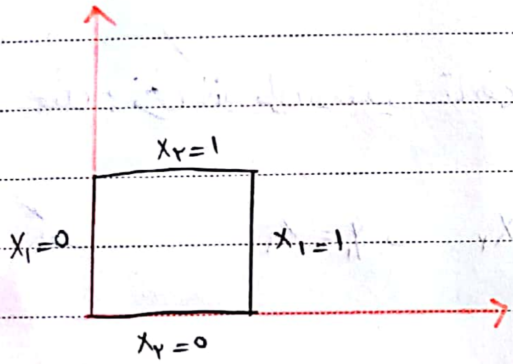
$$|J| = \begin{vmatrix} \frac{1}{2\sqrt{y_1}} & 0 \\ \frac{-1}{2}\frac{y_2}{y_1^{3/2}} & \frac{1}{\sqrt{y_1}} \end{vmatrix} = \frac{1}{2y_1}$$

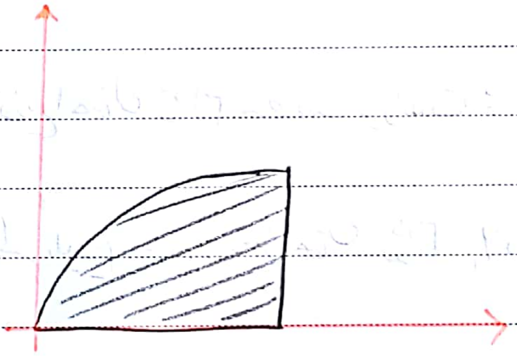
قدر مطلق دترمینان

۳) جایگذاری در $f_{X_1X_2}$

$$f_{Y_1Y_2} = \begin{cases} 4\sqrt{y_1} \cdot \frac{y_2}{\sqrt{y_1}} \cdot \frac{1}{2y_1} = \frac{2y_2}{y_1} \\ 0 \end{cases}$$

۴) مشخص کردن دامنه y_1 و y_2





$$\begin{cases} x_1=0 \rightarrow y_1 = x_1^2 \rightarrow y_1 = 0 \\ x_1 = 1 \rightarrow y_1 = x_1^2 \rightarrow y_1 = 1 \\ x_2 = 0 \rightarrow y_2 = x_1 x_2 \rightarrow y_2 = 0 \\ x_2 = 1 \rightarrow y_2 = x_1 x_2 \rightarrow y_2 = \sqrt{y_1} \end{cases}$$

مرزها را جایگذاری کنید در فرمول های y

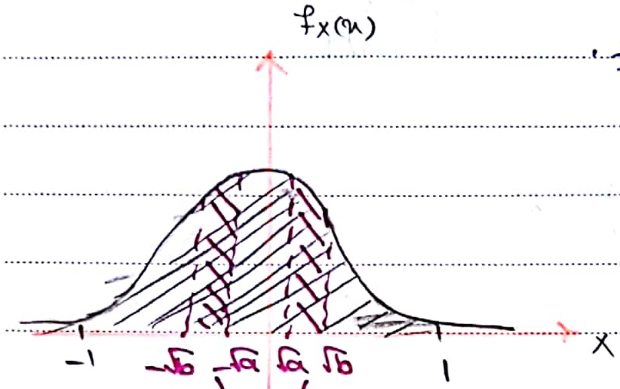
مرزهای جدید را بدست آورید.

$0 < y_1 < 1$

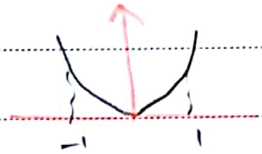
$0 < y_2 < \sqrt{y_1}$

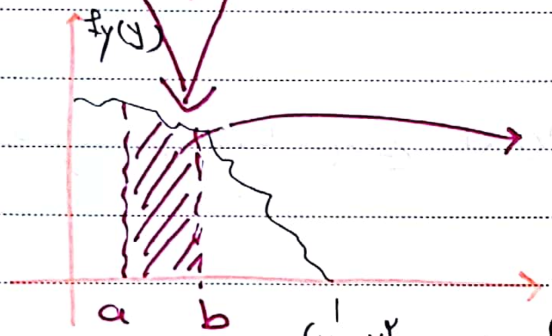
Subject:
Year. Month. Date. ( )

مثال: اگر تابع توزیع متغیر تصادفی X به صورت $f_X(x)$ پیوسته باشد، $y=x^2$ تعریف شود در دامنه X بازه $-1<X<1$ باشد آنگاه تابع توزیع y را بدست آورید.



$y=x^2$





$P(a<y<b)$

$(y=x^2 \to x=\sqrt{y}\ ,\ -\sqrt{y})$

$P(a<y<b) = P(-\sqrt{b}<x<-\sqrt{a}) + P(\sqrt{a}<x<\sqrt{b})$

$= \underbrace{\int_{-\sqrt{b}}^{-\sqrt{a}} f_X(x)\,dx}_{A} + \underbrace{\int_{\sqrt{a}}^{\sqrt{b}} f_X(x)\,dx}_{B}$

تغییر متغیر
$X=-\sqrt{y}$
$-1<X<0$

تغییر متغیر
$X=\sqrt{y}$
$0<X<1$

$\Rightarrow A+B = \int_b^a f_X(-\sqrt{y})\cdot\frac{-1}{2\sqrt{y}}\,dy + \int f_X(\sqrt{y})\cdot\frac{1}{2\sqrt{y}}\,dy$

$-1<X<0$
$X=-\sqrt{y}$
$dx=-\frac{1}{2\sqrt{y}}\,dy$

$0<X<1$
$X=\sqrt{y}$
$dx=\frac{1}{2\sqrt{y}}\,dy$

Subject:
Year. Month. Date. ( )

$$= \int_a^b f_X(-\sqrt{y}) \cdot \frac{1}{2\sqrt{y}}\, dy + \int_a^b f_X(\sqrt{y}) \frac{1}{2\sqrt{y}}\, dy$$

$$\left|\frac{\partial x}{\partial y}\right| \qquad \left|\frac{\partial x}{\partial y}\right|$$

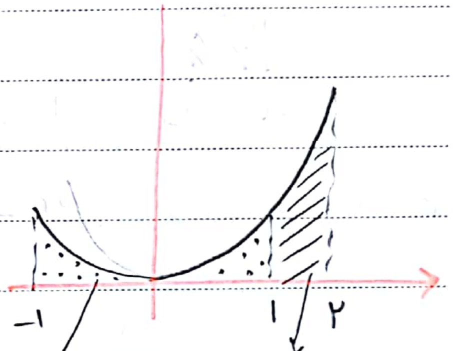
$$= \int_a^b f_X(-\sqrt{y}) + f_X(\sqrt{y}) \left(\left|\frac{\partial x}{\partial y}\right|\right) \cdot dy$$

$$\frac{1}{2\sqrt{y}}$$

$$= \int_a^b f_Y(y)\, dy \Longrightarrow f_Y(y) = \left(f_X(-\sqrt{y}) + f_X(\sqrt{y})\right) \frac{1}{2\sqrt{y}}$$

مثال: اگر متغیر تصادفی X دارای تابع توزیع $-1<x<2$ باشد آنگاه تابع توزیع $y=x^2$ را بدست آورید.

$$f_Y(y) = \begin{cases} f_X(-\sqrt{y}) + f_X(\sqrt{y}) & 0<y<1 \\ f(\sqrt{y}) \cdot \frac{1}{2\sqrt{y}} & 1<y<4 \\ 0 & \text{o.w} \end{cases}$$



قسمت غیر یک به یک جداگانه حل شود.

$y=x^2 \longrightarrow x=\sqrt{y}$

قسمت یک به یک جداگانه

* نکته: اگر متغیر تصادفی X دارای تابع توزیع $f_X(x)$ باشد و $y=u(x)$ در نظر گرفته شود، چنان u تابعی غیر یک به یک از X به Y باشد، X را به بازه هایی که تابع u روی آنها یک به یک باشد تقسیم می کنیم به طوری که معکوس تابع ها برای X بدون ابهام قابل محاسبه باشد یعنی داشته باشیم $X_1=w_1(y), X_2=w_2(y), \dots$

آنگاه: $X_n=w_n(y)$

$$f_Y(y) = \sum_{i=1}^{n} f_X(w_i(y)) \cdot |J_i|$$

که $J_i = \frac{dx_i}{dy}$

Subject:

Year. Month. Date. ( )

جلسه نوزدهم

آزمون های آماری (آزمون فرضیه):

نمونه های یک جمعیت یا دو جمعیت انجام می شود.

توزیع جمعیت:

$E[X]$ ✓ میانگین توزیع

$var[X]$ ✓

آماره: هر تابع یا معیاری که براساس نمونه ای از جمعیت محاسبه شود.

$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n}$: میانگین نمونه = آماره میانگین

$n \rightarrow$ سایز نمونه

$S^2 = \frac{\sum (x_i - \bar{X})^2}{n-1}$: واریانس نمونه = آماره واریانس

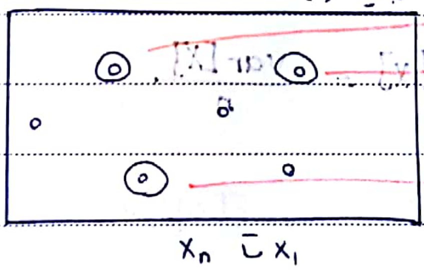
آماره میانگین $(\bar{X})$

$f_{\bar{X}}(\bar{X})$

$E[\bar{X}]$

$var[\bar{X}]$

جمعیت $f(x)$ — یک نمونه n تایی به تصادف n



X_1 یک متغیر تصادفی با توزیع $f(x)$ ← همان توزیع جمعیت

X_2 یک متغیر تصادفی با توزیع f، مستقل از بقیه متغیرها x_i

$\vdots$

x_n

اگر یک نمونه n تایی از جمعیت با توزیع $f(x)$ به تصادف انتخاب نمائیم ($x_1, \dots, x_n$) $X_1 \dots X_n$

هر کدام یک متغیر تصادفی با توزیع $f(x)$ خواهند بود و از یکدیگر مستقل در نظر گرفته می شوند. به عبارتی

$X_1 \dots X_n$، n متغیر تصادفی یکسان و متغیر خواهند بود که تابع توزیع احتمال توأم آنها به صورت زیر خواهد بود.

$$f_{x_1 \dots x_n}(x_1, x_2, \dots, x_n) = \prod_{i=1}^{n} f_{x_i}(x_i)$$

$$* \quad E[\bar{X}] = E[X] = \mu_x$$

اثبات $\longrightarrow$ $E[\bar{X}] = E\left[\dfrac{X_1 + X_2 + X_3 + \dots + X_n}{n}\right]$

$$= \frac{1}{n}\left(E[X_1] + E[X_2] + \dots + E[X_n]\right) = \frac{n}{n} E[X]$$

$$E[X_1] = E[X_2] = \dots E[X_n]$$

$$Var[\bar{X}] = var\left[\frac{X_1 + X_2 + \dots + X_n}{n}\right] = \frac{1}{n^2}\left(\underset{=}{var[X_1]} + \dots + \underset{=}{var[X_n]}\right)$$

$$= \frac{n}{n^2} var[X] = \frac{var[X]}{n}$$

* اگر جمعیت دارای توزیع نرمال باشد، میانگین $E[X] = \mu_x$ و $var[X] = \sigma_x^2$ ، آنگاه میانگین

نمونه های n تایی از این جمعیت دارای توزیع نرمال با پارامترهای $\mu_{\bar{x}} = \mu_x$ و $\sigma_{\bar{x}}^2 = \dfrac{\sigma_x^2}{n}$

است.

قضیه حد مرکزی central Limit Theoreme :

(در مورد تابع توزیع $\bar{x}$ صحبت می‌کند.)

اگر $\bar{x}$ میانگین یک نمونه تصادفی با سایز n از یک توزیع (جمعیت) یا (متغیر ریاضی) میانگین μ_x و واریانس σ_x^2 باشد، نگاه اگر $n \to \infty$ تابع توزیع احتمال $\bar{x}$ توزیع نرمال با میانگین $E[\bar{X}] = \mu_x$

و $var[\bar{X}] = \frac{\sigma_x^2}{n}$ خواهد بود.

$n \geq 30 \longrightarrow$ تابع توزیع $\bar{x}$ یک متغیر نرمال است

$n < 30 \longrightarrow$ خود داده‌ها نرمال باید نرمال باشند $\longrightarrow$ باز هم نرمال است.

اگر $n \geq 30$ باشد قضیه حد مرکزی برقرار است. اگر $n < 40$ باشد و جمعیت دارای توزیع بسیار نرمال باشد هم قضیه حد مرکزی قابل استفاده است.

مثال: طول عمر یک قطعه الکترونیکی به سبب کارخانه تولیدی بشد دارای توزیع نرمال است و میانگین طول عمر این قطعه ۸۰۰ است و انحراف معیار ۴۰ ساعت. اگر یک نمونه ۱۶ تایی از آنها به تصادف انتخاب کنیم چقدر احتمال دارد میانگین طول عمر آنها از ۷۷۵ ساعت بیشتر باشد؟

$$\int_{775}^{+\infty} N(800, 10) = 1 - CDF\left(\frac{775-800}{10}\right) = 1 - CDF(-2.5)$$

توزیع نرمال استاندارد

$1 - 0.0062$

$$x^* = \bar{X}_1 - \bar{X}_2$$

$$E[x^*] = E[\bar{X}_1] - E[\bar{X}_2]$$

$$var[x^*] = var[\bar{X}_1] - var[\bar{X}_2] = \frac{\sigma_{x_1}^2}{n_1} + \frac{\sigma_{x_2}^2}{n_2}$$

۹۷,۳,۱

جلسه بیستم

آزمون فرضیه:

توزیع اختلاف میانگین های نمونه از دو جمعیت:

اگر از دو جمعیت (توزیع) نمونه به تصادف انتخاب نماییم به طوری که اندازه نمونه جمعیت اول n_1 و اندازه نمونه جمعیت دوم n_2 باشد. در صورتی که $n_1, n_2 \geq 30$ باشد و یا توزیع دو جمعیت نرمال باشد آنگاه متغیر تصادفی $\bar{X}_1 - \bar{X}_2$ دارای توزیع نرمال با میانگین $\mu_{x_1} - \mu_{x_2}$ و انحراف معیار $\sqrt{\frac{\sigma_{x_1}^2}{n_1} + \frac{\sigma_{x_2}^2}{n_2}}$ است که در آن μ_1 و μ_2 به ترتیب میانگین جمعیت اول و دوم و σ_1 و σ_2 انحراف معیار جمعیت اول و دوم است.

قضیه حد مرکزی: $n \to \infty$

یا $n \geq 30$

میانگین

$$Z = \frac{\bar{X} - \mu_x}{\frac{\sigma_x}{\sqrt{n}}}$$

Z دارای توزیع نرمال استاندارد است.

آماره

$E[\bar{X}] = \mu_x$

sample mean

میانگین نمونه

انحراف معیار

$var[\bar{X}] = \frac{\sigma^2}{n}$

سایز نمونه

$\bar{X} \sim N\left(\mu_x, \frac{\sigma_x^2}{n}\right)$

واریانس

Subject:
Year. Month. Date. ( )

$$z = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_{x_1} - \mu_{x_2})}{\sqrt{\frac{\sigma_{x_1}^2}{n_1} + \frac{\sigma_{x_2}^2}{n_2}}}$$

می توان گفت متغیر تصادفی z دارای توزیع نرمال استاندارد است.

مثال: قطعات الکترونیکی تولید یک کارخانه دارای طول عمر $\underbrace{4{,}5}_{\mu_{x_1}}$ سال (میانگین) و انحراف معیار $\underbrace{0{,}9}_{\sigma_{x_1}}$ سال است و قطعات مشابهی که در کارخانه دیگری تولید می شود دارای طول عمر متوسط $\underbrace{4}_{\mu_{x_2}}$ سال و انحراف معیار $\underbrace{0{,}8}_{\sigma_{x_2}}$ سال است اگر یک نمونه $\underbrace{36}_{n_1}$ تایی از قطعات کارخانه اول و یک نمونه $\underbrace{49}_{n_2}$ تایی از قطعات کارخانه دوم انتخاب نماییم. چقدر احتمال دارد میانگین طول عمر نمونه کارخانه اول بیش از یک سال بیشتر از میانگین طول عمر قطعات نمونه کارخانه دوم باشد؟

$$\underbrace{P(\bar{X}_1 - \bar{X}_2 > 1)}_{\text{توزیع نرمال}} = \int_1^{+\infty} N\left(\underbrace{4{,}5 - 4}_{\mu_{x_1} - \mu_{x_2}}, \sqrt{\frac{(0{,}9)^2}{36} + \frac{(0{,}8)^2}{49}}\right)$$

$$= 1 - CDF\left(\frac{\overbrace{\bar{X}_1 - \bar{X}_2}^{1} - (0{,}5)}{\sqrt{\frac{(0{,}9)^2}{36} + \frac{(0{,}8)^2}{49}}}\right) = 1 - \underbrace{CDF(2{,}65)}_{0{,}9945} \simeq 0{,}01$$

* اگر اندازه نمونه کوچکتر از 30 بود یا اینکه برای جمعیت انحراف معیار را نداشته باشیم برای تقریب توزیع $\bar{X}$ از توزیع t استفاده می کنیم.

آزمون فرضیه (تست آماری): تعریف

روشی است که طی آن با استفاده از روش‌های آماری می‌خواهیم درستی یا نادرستی یک ادعا را بیازماییم

فرضیه: ادعای است درباره یک یا چند جمعیت که ممکن است درست یا غلط باشد.

مثال ①

مثال

میانگین زمان اجرای الگوریتم ① از الگوریتم ② کمتر است. $\bar{X}_1 - \bar{X}_2$

میانگین زمان اجرای الگوریتم ① برابر ۵۰۰ است. $\bar{X}_1 = 500$

شاخص (آماره) آزمون:

شاخصی است که مقدار آن را به کمک یک نمونه تصادفی بدست می‌آوریم و به عنوان معیاری برای رد یا قبول فرضیه

به کار می‌بریم.

مثال ①: $\mu_{x_1} = \mu_{x_2}$ جمعیت

صفر

H_0: فرضیه‌ای است که همواره شامل تساوی است و اغلب هدف رد آن است.

H_1: فرضیه مقابل (یا جایگزین) یا فرضیه‌ای است که در مقابل فرضیه صفر (H_0) مطرح می‌شود و در صورت

شامل نامساوی

رد H_0 آن را می‌پذیریم. $\mu_{x_1} < \mu_{x_2}$

سطح معناداری (α):

میزان خطایی که ممکن است در رد فرضیه صفر (H_0) مرتکب شویم. به عبارتی احتمال رد کردن فرضیه H_0 وقتی

H_0 درست است، که به آن خطای نوع اول هم می‌گوییم.

	فرض H_0 درست است	فرض H_0 نادرست است
رد فرض H_0	خطای نوع اول (سطح معناداری) α	قدرت آزمون $1-\beta$
قبول فرض H_0	سطح اطمینان $1-\alpha$	خطای نوع دوم β

۹۷/۳/۶

جلسه بیست دوم

مراحل آزمون فرضیه (روش اول) "آزمون فرضیه میانگین"

۱) تشکیل دادن فرضیه های آزمون

۲) مشخص کردن یک سطح معناداری

۳) انتخاب آماره آزمون

۴) مشخص کردن ناحیه بحرانی براساس α (مشخص کردن ناحیه رد یا قبول فرض H_0)

۵) رد H_0 درصورتی که آماره آزمون در ناحیه بحرانی (رد) واقع گردد.

۶)
اگر H_0 رد شد: مشاهدات کافی برای قبول H_0 وجود ندارد
اگر H_0 قبول شد: مشاهدات کافی برای رد H_0 وجود ندارد.

مثال: یک کارخانه تولید چوب ماهیگیری ادعا دارد چوب های تولیدی آن دارای میانگین تحمل بار 8kg با انحراف معیار 0/5kg است. یک مسئول کنترل کیفیت برای یک نمونه ۵۰ تایی میانگین تحمل بار را برابر 7/8kg بدست آورده. صحت ادعای کارخانه را با یک آزمون فرضیه با $\alpha = 0/01$ بررسی نمائید

الف) با یک آزمون فرضیه یک طرفه ب) با یک آزمون فرضیه دو طرفه

Subject: ۸۹

Year. Month. Date. ( )

① $\left\{ \begin{array}{l} H_0 : \mu_x = 8 \\ H_1 : \underbrace{\mu_x < 8\,kg}_{\text{آزمون فرضیه یک طرفه}} \qquad \boxed{\underbrace{\mu_x \neq 8\,kg}_{\text{آزمون فرضیه دو طرفه}}} \end{array} \right.$

* با توجه به اطلاعات نمونه هدف رد یا قبول فرضیه ای درباره جمعیت است.

برای آزمون فرضیه یک طرفه (الف) ***

② $\alpha = 0.01$

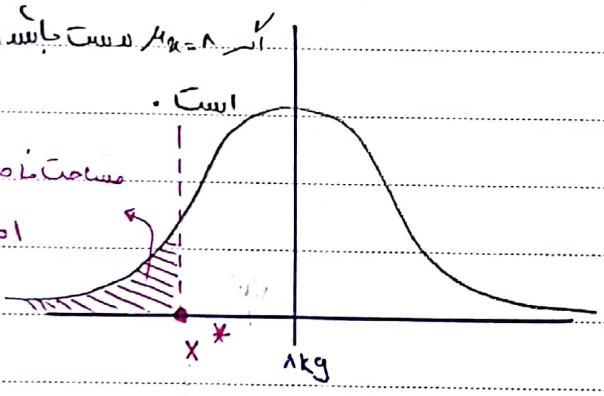
* فرضیه حتماً در مورد جمعیت و نه نمونه

* H_0 شامل تساوی است.

③ آماره آزمون : $\bar{X} \xrightarrow{\text{به خاطر توزیع نرمال استاندارد}} \underbrace{Z}_{\text{آماره آزمون}} = \dfrac{X - \mu_x}{\frac{\sigma_x}{\sqrt{n}}}$

اگر $\mu_x = 8$ درست باشد توزیع $\bar{X}$ به صورت زیر است.

④ مساحت ناحیه برای $\alpha = 0.01$



$$CDF\left(\frac{x^* - 8}{\frac{0.5}{\sqrt{50}}}\right) = 0.01$$

$$\frac{x^* - 8}{\frac{0.5}{\sqrt{50}}} = -2.32$$

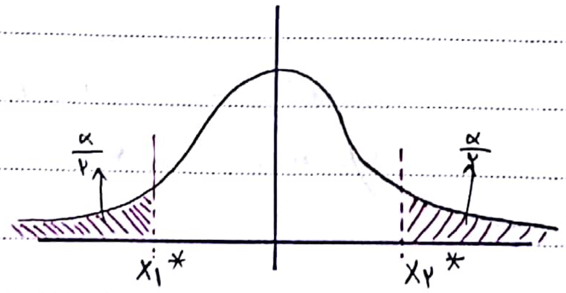
$$\Rightarrow \boxed{x^* = 7.83}$$

$\Leftarrow$ با توجه به اینکه ۷٫۸ در ناحیه رد H_0 قرار دارد

برای آزمون فرضیه دو طرفه (ب)

$\left\{ \begin{array}{ll} H_0 & \mu_x = 8 \\ H_1 & \mu_x \neq 8 \end{array} \right.$

* در آزمون دو طرفه به ازای α برابر با آزمون یک طرفه رد کردن H_0 سخت است. بنابراین آزمون دقیق تری است.

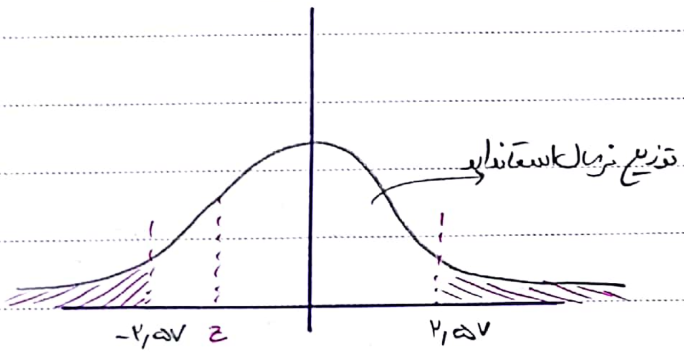


X_1^* : $CDF\left(\frac{X_1^*-8}{\frac{0.5}{\sqrt{50}}}\right)=0.005$

$$\frac{X_1^*-8}{\frac{0.5}{\sqrt{50}}}=-2.57 \Rightarrow \boxed{X_1^*=7.82}$$

X_2^* : $1-CDF\left(\frac{X_2^*-8}{\frac{0.5}{\sqrt{50}}}\right)=0.005$

$$\frac{X_2^*-8}{\frac{0.5}{\sqrt{50}}}=2.57$$



$$\frac{\bar{X}-\mu_x}{\frac{\sigma_x}{\sqrt{n}}}=\frac{\bar{X}-8}{\frac{0.5}{\sqrt{50}}}$$

$$z^*=\frac{7.8-8}{\frac{0.5}{\sqrt{50}}}$$

* حل: H_0 را رد کرده ایم با احتمال α ممکن است H_0 درست باشد و به اشتباه آن را رد کرده باشیم

توزیع t :

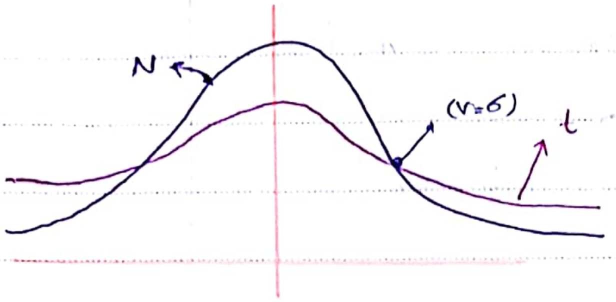
اگر برای یک جمعیت انحراف معیار را نداشته باشیم یا $n<30$ باشد آنگاه $\frac{\bar{X}-\mu_x}{\frac{s}{\sqrt{n}}}$ دارای توزیع t

انحراف معیار نمونه

با $n-1$ درجه آزادی است

$v=n-1$

فرق توزیع t با توزیع نرمال :



جلسه بیست دوم

۹۷ ,۳ ,۸

آزمون فرضیه میانگین :

① آزمون فرضیه درمورد میانگین یک جمعیت :

① تعداد نمونه بیش از ۳۰ است و انحراف معیار را می دانیم

← آماره آزمون $z=\frac{\bar{x}-\mu_x}{\frac{\sigma_x}{\sqrt{n}}}$ ، z دارای توزیع نرمال استاندارد است .

② تعداد نمونه کمتر از ۳۰ است یا انحراف معیار جمعیت را نمی دانیم

آماره آزمون $t=\frac{\bar{x}-\mu_x}{\frac{S}{\sqrt{n}}}$ دارای توزیع t با $v=n-1$ درجه آزادی است

انحراف معیار نمونه

$$\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}$$

② آزمون فرضیه درمورد میانگین دو جمعیت :

① تعداد نمونه برای هر دو جمعیت بیش از ۳۰ است و انحراف معیار هر دو جمعیت را می دانیم .

$H_0: \mu_{x_1}-\mu_{x_2}=d$

← آماره آزمون

$$z=\frac{(\bar{X}_1-\bar{X}_2)-\overbrace{(\mu_{x_1}-\mu_{x_2})}^{d}}{\sqrt{\frac{\sigma_{x_1}^2}{n_1}+\frac{\sigma_{x_2}^2}{n_2}}}$$

دارای توزیع نرمال استاندارد است .

$\bar{X}_1-\bar{X}_2$ دارای توزیع نرمال

$$E[\bar{X}_1-\bar{X}_2] \rightarrow \mu_{x_1}-\mu_{x_2}$$

$$Var[\bar{X}_1-\bar{X}_2] \rightarrow \frac{\sigma_{x_1}^2}{n_1}+\frac{\sigma_{x_2}^2}{n_2}$$

$$N\left(\mu_{x_1}-\mu_{x_2}, \sqrt{\frac{\sigma_{x_1}^2}{n_1}+\frac{\sigma_{x_2}^2}{n_2}}\right)$$

توزیع $\bar{X}_1-\bar{X}_2$

۲) اگر n ها کوچکتر از ۳۰ بود یا انحراف معیار جمعیت را نداشتیم.

الف) می‌توان فرض کرد $\sigma_{x_2} = \sigma_{x_1}$

آماره آزمون

$$t = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_{x_1} - \mu_{x_2})}{\sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}}$$

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

متغیر تصادفی t دارای توزیع t با $v = n_1 + n_2 - 2$ درجه آزادی است.

ب) اگر نتوانیم فرض کنیم $\sigma_{x_1} = \sigma_{x_2}$

آماره آزمون

$$t = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_{x_1} - \mu_{x_2})}{\sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}}$$

دارای توزیع t با v درجه آزادی

$$v = \frac{\left(S_1^2/n + S_2^2/n\right)^2}{\frac{(S_1^2/n)^2}{n_1 - 1} + \frac{(S_2^2/n)^2}{n_2 - 1}}$$

(ttest) (مثل های اسلایدها)

Subject

Date

۹۷,۳,۱۳

جلسه بیست و ششم (جلسه آخر)

خطا و دقت آزمون فرضیه

	H_0 درست باشد	H_0 نادرست باشد
H_0 رد شود	خطای نوع اول یا سطح معناداری احتمال = α	سطح اطمینان احتمال = $1-\alpha$
H_0 قبول شود	دقت آزمون احتمال = $1-\beta$	خطای نوع دوم احتمال = β

α: احتمال اینکه فرض H_0 درست باشد و به اشتباه رد شود.

β: احتمال اینکه فرض H_0 نادرست باشد و به اشتباه قبول شود

* همواره خطای نوع I برابر مساحت ناحیه رد (ناحیه بحرانی) است

مثال: (10.15)

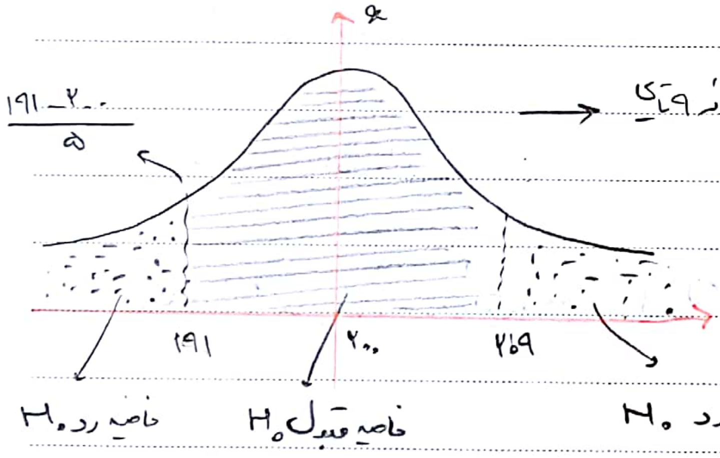
$H_0 \quad \mu = 200 \text{ mL}$

$H_1 \quad \mu \neq 200$

$n = ۹$

$۱۹۱ < \bar{x} < ۲۰۹$

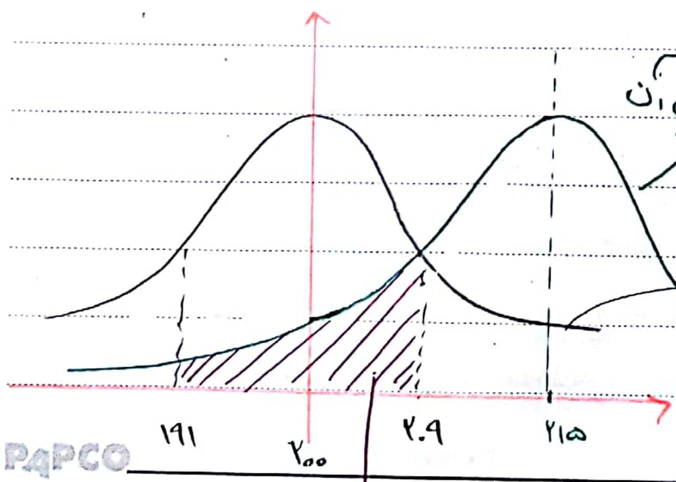
خطای نوع اول: میانگین $\mu = ۲۰۰$ باشد ما به اشتباه آن را رد کنیم



اگر میانگین جامعه واقعاً ۲۰۰ باشد توزیع $\bar{x}$ برای نمونه ۹ تایی به صورت روبرو است

$E[\bar{x}] = ۲۰۰$

$\sigma[\bar{x}] = \frac{\sigma_x}{\sqrt{۹}} = \frac{۱۵}{۳} = ۵$



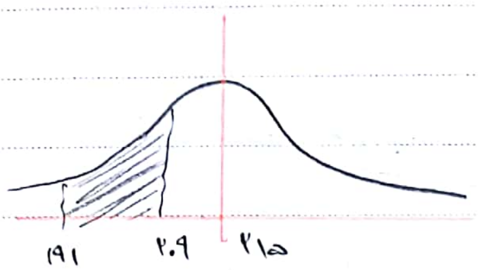
تابع توزیع $\bar{x}$ برای نمونه‌های ۹ تایی از جامعه‌ای به میانگین $\mu = ۲۱۵$ باشد

تابع توزیع $\bar{x}$ برای نمونه‌های ۹ تایی از جامعه‌ای که میانگین آن $\mu = ۲۰۰$ باشد.

این مساحت = احتمال خطای نوع دوم

احتمال خطای نوع دوم: اگر میانگین جامعه واقعاً ۲۱۵ باشد چقدر احتمال دارد تا میانگین یک نمونه ۹ تایی از آن

در ناحیه قبول H_0 یا بازه $191 < \bar{x} < 209$ بیفتد؟



$$E[\bar{X}] = 215$$

$$\sigma_{\bar{X}} = \frac{15}{3} = 5$$

مثال: یک کارخانه ادعا دارد چوب‌های ماهیگیری تولیدی آن دارای میانگین تحمل وزن $\mu = 15kg$ و انحراف معیار

$\sigma = 0.5 kg$ است. برای ارزیابی صحت ادعای آن آزمون فرضیه‌ای با فرض $H_0 : \mu = 15kg$

و فرض $H_1 : \mu < 15$ با یک نمونه ۵۰ تایی انجام می‌دهیم و اگر برای نمونه $\bar{X} > 14.9$ شد

ادعای کارخانه را می‌پذیریم.

الف) احتمال خطای نوع اول را بیابید.

ب) احتمال خطای نوع دوم را اگر میانگین واقعی تحمل وزن برابر $\mu = 14.8$ و $\mu = 14.9$ باشد بیابید.

$$\begin{cases} H_0 & \mu = 15 \\ H_1 & \mu < 15 \end{cases}$$

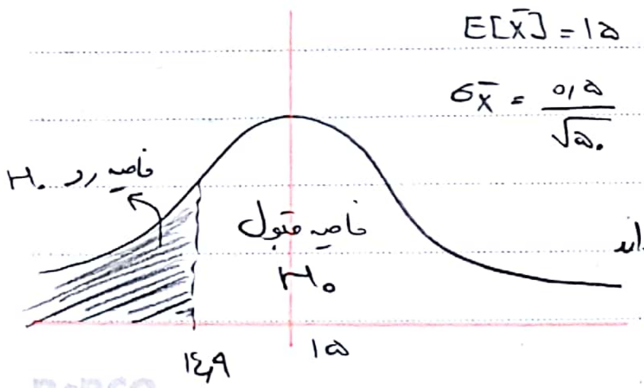
$n = 50$

به جای α ناحیه رد یا قبول مشخص شده

α : مساحت ناحیه رد یا بحرانی

$$E[\bar{X}] = 15$$

$$\sigma_{\bar{X}} = \frac{0.5}{\sqrt{50}}$$

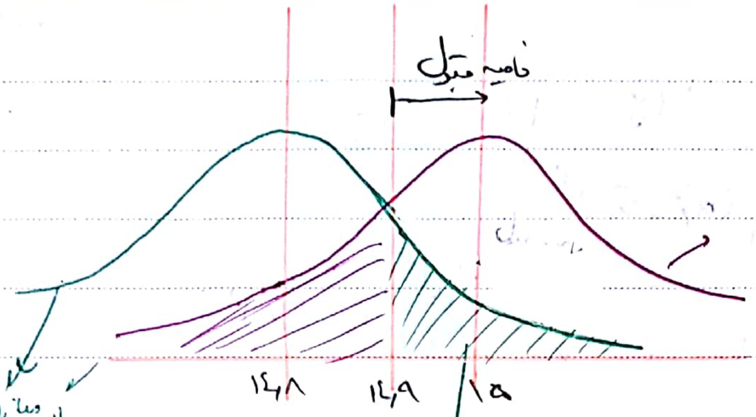


$$\alpha = CDF\left(\frac{14.9 - 15}{\frac{0.5}{\sqrt{50}}}\right) = 0.07$$

توزیع نرمال استاندارد

حالا احتمال ۰٫۰۷ ممکن است کارخانه واقعاً $\mu = 15$ داشته باشد

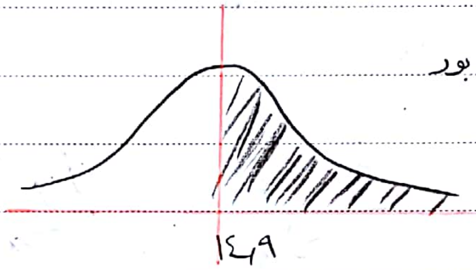
ولی یک نمونه ۵۰ تایی $\bar{X} < 14.9$ شود.



اگر میانگین واقعاً ۱۵ بود و توزیع $\bar{X}=15$

اگر میانگین جامعه به جای
$\mu=15$ ، $\mu=14.8$ بود
آنگاه توزیع $\bar{X}$
$E[\bar{X}]=14.8$
$\sigma_{\bar{X}} = \frac{0.5}{\sqrt{50}}$
انحراف معیار را مشابه قبلی حساب می‌کنیم

مساحت یا احتمال خطای نوع دوم $= \beta$



اگر میانگین واقعاً ۱۴٫۹ بود
توزیع $\bar{X}$
$\sigma_{\bar{X}} = \frac{0.5}{\sqrt{50}}$

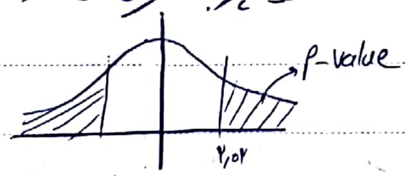
مراحل آزمون فرضیه روش دوم (در امتحان نمی‌آید)

۱) مشخص کردن فرضیه H_0 و H_1

۲) مشخص کردن آماره آزمون

۳) محاسبه p-value

۴) انجام آزمون براساس p-value : اگر p-value از یک حد اساسی (α) کوچکتر بود فرض H_0 رد می‌شود.



p-value کوچکترین αای است که می‌توان با آن فرض H_0 را رد کرد.
احتمال خطای نوع اول

جزو درس می باشد
جمله سؤال نمی آید

$$b_1 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \quad \text{I}$$

sample variance واریانس نمونه

$$= S_x^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$$

sample covariance کواریانس نمونه

$$S_{xy} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{n-1}$$

از I داریم:

$$\frac{(n-1)\, S_{xy}}{(n-1)\, S_x} = \frac{\rho_{xy} S_x S_y}{S_x^2} = \frac{\rho_{xy} S_y}{S_x}$$

ρ_{xy} در شیب خط رگرسیون ظاهر می شود.

$$\rho_{xy} = \frac{\mathrm{cov}(x,y)}{\sigma_x \sigma_y}$$

* علامت شیب خط را ضریب همبستگی مشخص می کند.