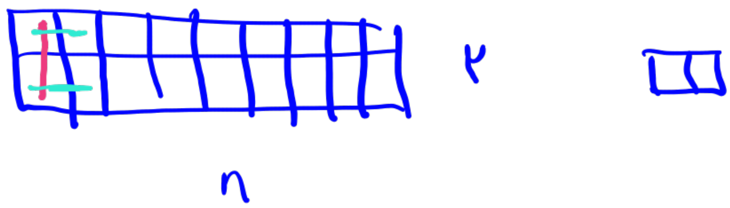
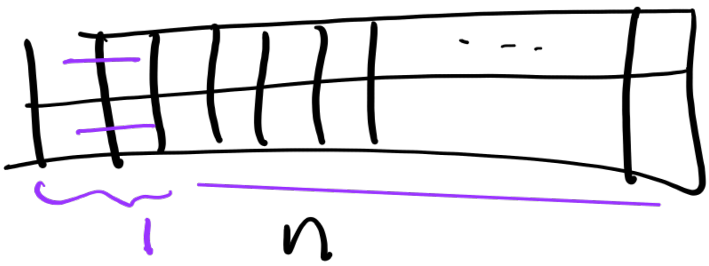


$$f_n = f_{n-1} + f_n$$



$$F_n = 1 \times F_{n-1} + 1 \times F_{n-2}$$

$$f_1 = 1 \qquad f_2 = 2 \qquad f_{n \geq 3} = f_{n-1} + f_{n-2}$$

$$f_0 = 0 \qquad f_1 = 1 \qquad f_2 = 1 \qquad f_3 = 2 \qquad \dots$$

$$F = \sum_{n=0}^{\infty} f_n x^n = f_0 + f_1 x + f_2 x^2 + \dots$$

$$0 + x + x^2 + 2x^3 + \dots$$

$$n \geq 2 \rightarrow f_n x^n = f_{n-1} x^n + f_{n-2} x^n$$

$$\sum_{n \geq 2} f_n x^n = x \sum_{n \geq 2} f_{n-1} x^{n-1} + x^2 \sum_{n \geq 2} f_{n-2} x^{n-2}$$

$$F - x = xF + x^2 F \Rightarrow F(1 - x - x^2) = x$$

$$F = \frac{x}{1 - x - x^2}$$

$\Omega = \{1, 2, \ldots, 6\}$ → Sample Space

فضای نمونه

2 → Observation مشاهده

$A = \{2, 4, 6\}$ پیشامد A واقع رخ داده است

$A \subset \Omega$

Event

پیشامد واقعه

اصل ۱: $P(A) \geq 0$

اصل ۲: $P(\Omega) = 1$

اصل ۳: $B \cap A = \phi \Rightarrow P(A \cup B) = P(A) + P(B)$

$P(\phi) =$

$$\phi \cap \Omega = \phi \to \overset{1}{P(\phi \cup \Omega)} = \overset{0}{P(\phi)} + \overset{1}{P(\Omega)}$$

$P(\bar{A}) \Rightarrow$

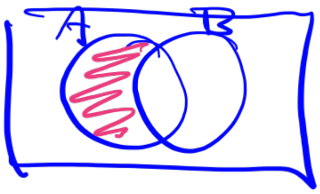
$A \cap \bar{A} = \phi$

$$\underbrace{P(A \cup \bar{A})}_{P(\Omega)} = P(A) + P(\bar{A})$$

$$1 = P(A) + P(\bar{A}) \to P(\bar{A}) = 1 - P(A)$$

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

$$A \cup B = (A-B) \cup B$$



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

$$P(A-B) =$$

$$A - B = A - (A \cap B) \rightarrow A = (A-B) \cup (A \cap B)$$

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

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

فرض کنید یک تست سرطان اختراع کرده‌ایم که

Sensitivity = ۹۹٪

Specificity = ۹۸٪

	نتیجه تست: منفی	نتیجه تست: مثبت
سرطان دارد	FN	TP
ندارد	TN	FP

$$\text{Sensitivity} = \frac{TP}{TP+FN}$$

$$\text{Specificity} = \frac{TN}{TN+FP}$$

در یک شهر احتمال ابتلا به سرطان 0/0001 است. اگر نتیجه تست برای یک نفر مثبت باشد چقدر احتمال دارد که او بیمار باشد؟

A : فرد سرطان دارد — B : نتیجه تست مثبت است

$\bar{A}$: ندارد — $\bar{B}$: منفی

$$P(A|B)$$

$$P(B|A) = \frac{99}{100}$$

$$P(\bar{B}|\bar{A}) = \frac{98}{100}$$

$$P(A) = \frac{1}{10000}$$

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

$$P(B) = P(B|A)P(A) + P(B|\bar{A})P(\bar{A})$$

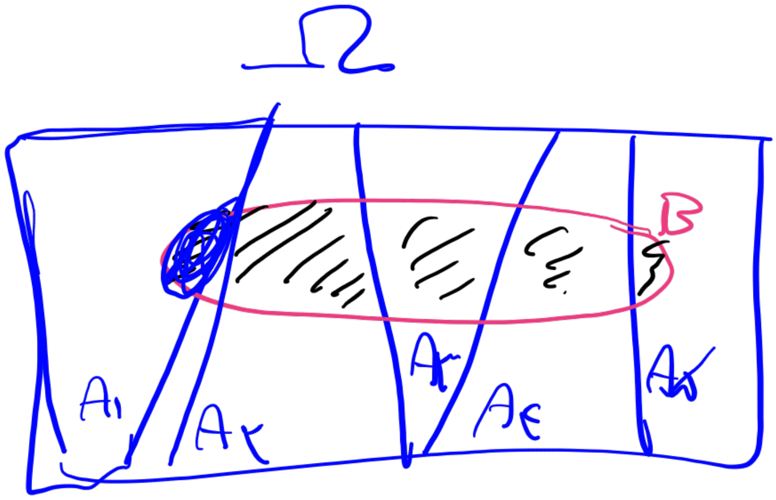
$$= \frac{\frac{99}{100} \times \frac{1}{10000}}{\frac{99}{10^4} + \frac{2}{100} \times \frac{9999}{10000}} = \frac{99}{99+}$$

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

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

Bayes فرمول

قانون احتمال کل :



$$\forall_{i \neq j} : \quad A_i \cap A_j = \phi$$

$$\bigcup_i A_i = \Omega$$

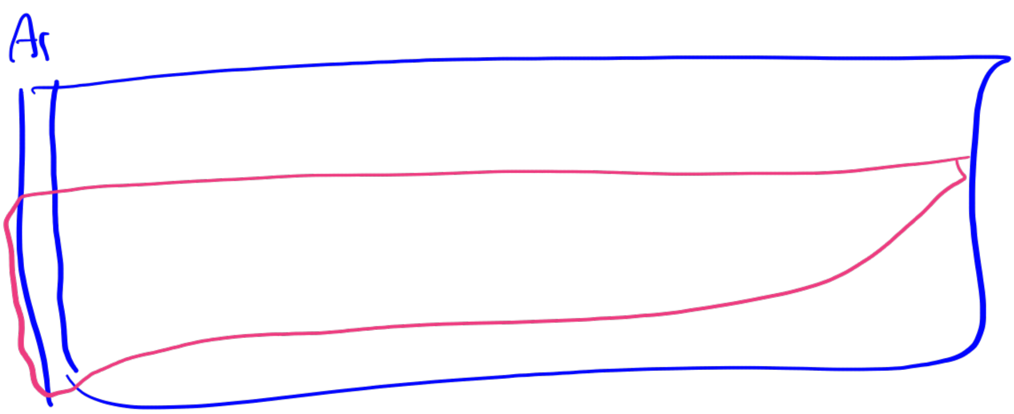
$$P(B) = P\big((B \cap A_1) \cup (B \cap A_2) \cup (B \cap A_3) \cup (B \cap A_4) \cup (B \cap A_5)\big)$$

$$= P(B \cap A_1) + P(B \cap A_2) + \cdots + P(B \cap A_5)$$

$$= \sum_i P(B \mid A_i)\, P(A_i)$$

$$P(B \cap A_1) = P(B \mid A_1)\, P(A_1)$$

$$P(B \mid A_1) = \frac{P(B \cap A_1)}{P(A_1)}$$



دو پیشامد A و B مستقل هستند اگر
iff
اطلاع از روی دادن B اثری در احتمال وقوع A ندارد.

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

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

اگر A و B مستقل باشند، $\bar{A}$ و B مستقل

$$P(\bar{A}|B) = 1 - P(A|B) = 1 - P(A) = P(\bar{A})$$

مستقل $\begin{cases} A \text{ و } \bar{B} \\ \bar{A} \text{ و } \bar{B} \end{cases}$

A و B و C مستقل هستند اگر:

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

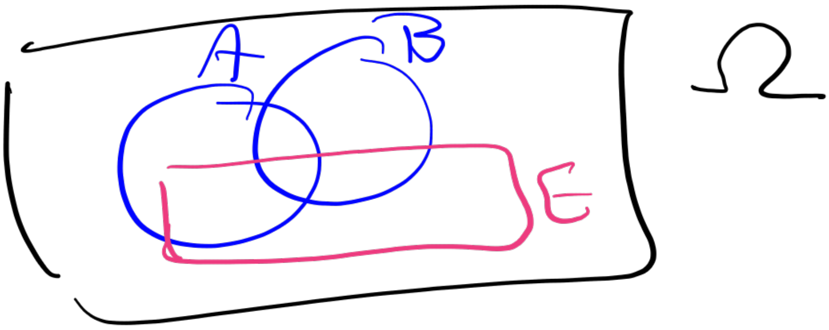
$$\left.\begin{aligned} P(A\cap B) &= P(A)P(B) \\ P(B\cap C) &= P(B)P(C) \\ P(A\cap C) &= P(A)P(C) \\ P(A\cap B\cap C) &= P(A)P(B)P(C) \end{aligned}\right\} \text{هم‌ارز}$$

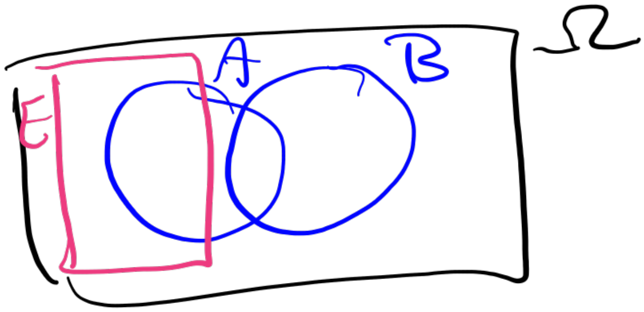
$A_1 \dots A_n$ مستقل هستند اگر و تنها اگر:

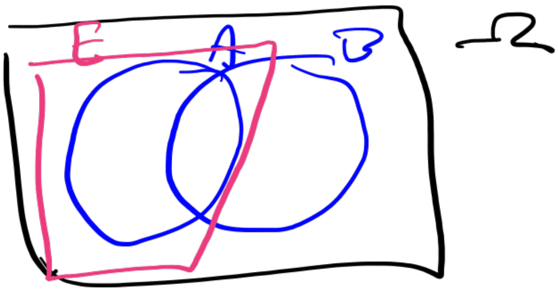
$$\forall\, k \subset \{1,\dots,n\} \quad P\Big(\bigcap_{j\in k} A_j\Big) = \prod_{j\in k} P(A_j)$$

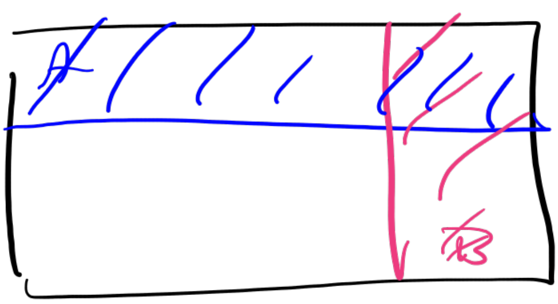
A و B با دانستن E به صورت شرطی مستقل هستند اگر

$$P(A \cap B \mid E) = P(A \mid E)\,P(B \mid E)$$

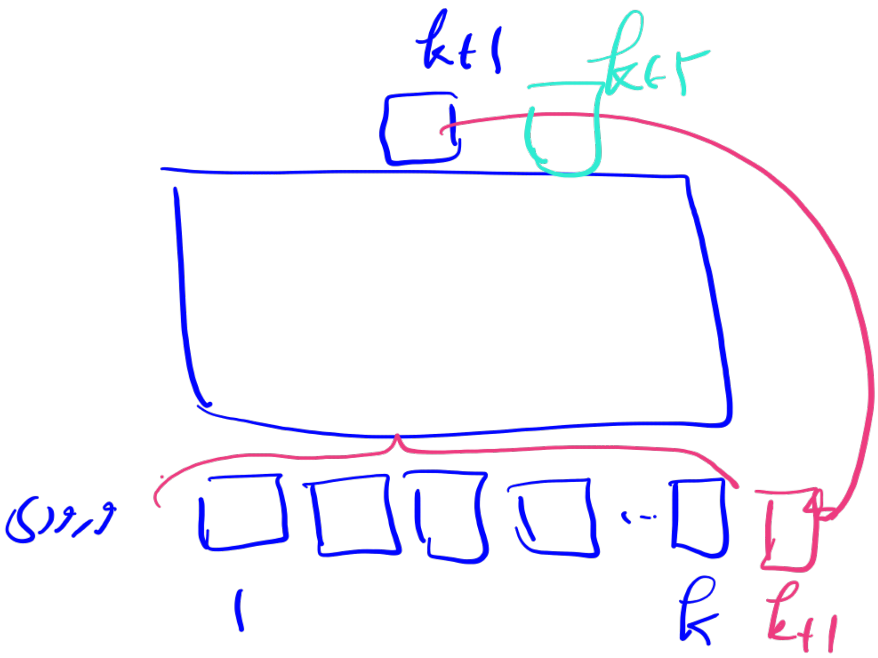






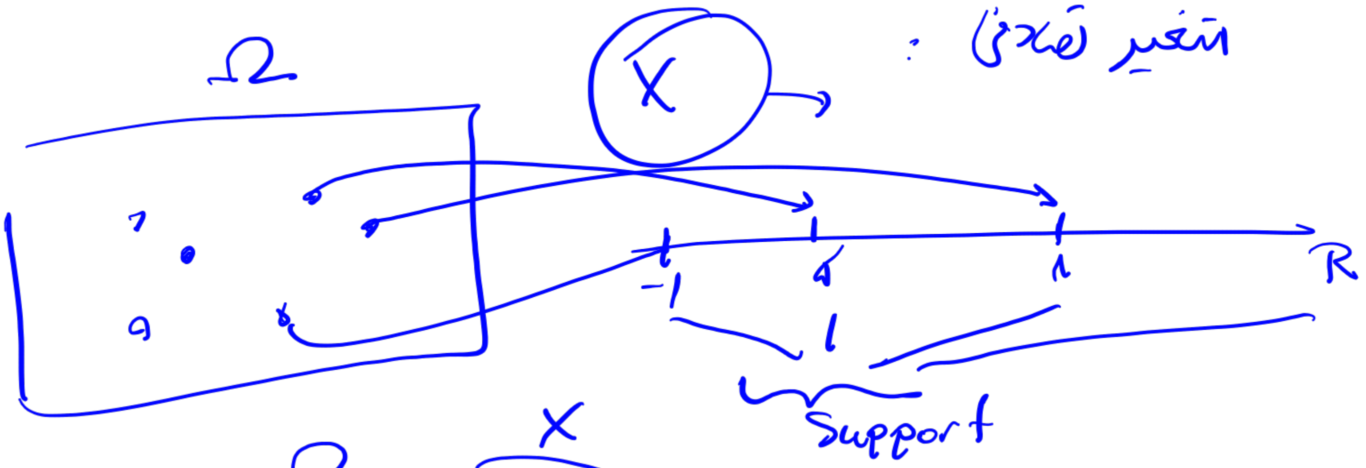


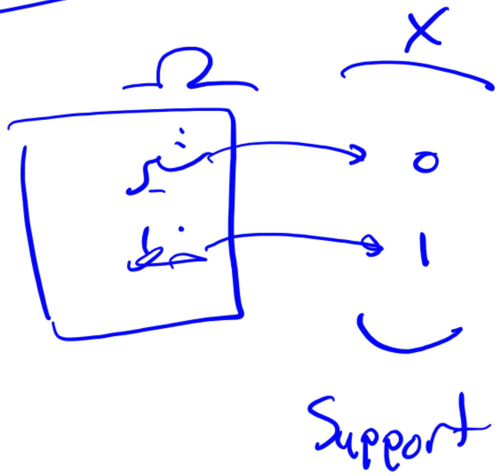
LLM Large Language Model



$$P(X \mid x_1, x_2, \ldots, x_k)$$

متغیر تصادفی :





$$Y = X + X$$

$$Y \underbrace{\begin{cases} 0 \\ 1 \\ ۲ \end{cases}}_{\text{Support}}$$

$$\underline{Y = ۲X}$$

متغیر تصادفی گسسته :

Support محدود

شمارا $\sim \mathbb{N}$

داشته باشد

تعداد فرزند ←

سن بر حسب سال

پیوسته : ناشمارا Support $\sim \mathbb{R}$

قد ←

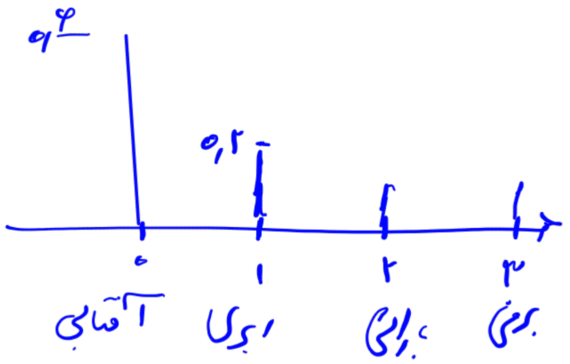
دما

طول عمر

Probability Mass Function   PMF

$E(X) = 0.7$

$f_X(x)$



X متغیر تصادفی

$P(X=0)=0.6 \quad P(X=2)=0.1$

$P(X=1)=0.2 \quad P(X=3)=0.1$

$\rightarrow f_X(0)=0.6 \quad f_X(2)=0.1$

$f_X(1)=0.2 \quad f_X(3)=0.1$

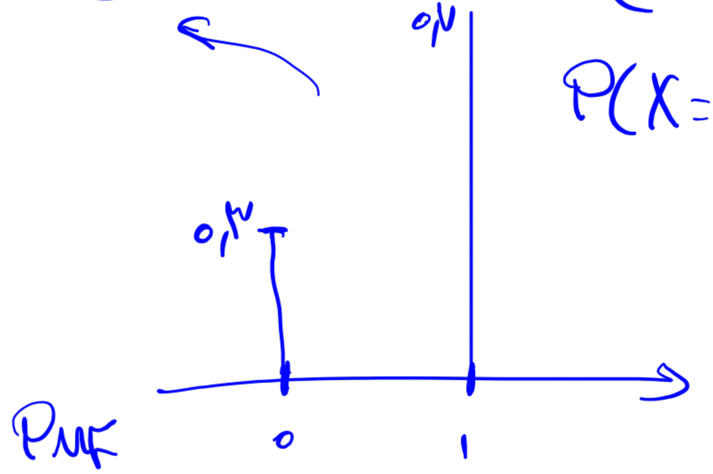
برنولی: انداختن سکه که با احتمال P برابر یک می‌آید و $1-P$ برابر صفر

$X \sim Bern(0.7)$

پیروی می‌کند

$P(X=0)= 0.3$

$P(X=1)= 0.7$



$f_X(0) = 0.3$

$f_X(1) = 0.7$

PMF:

- در شهرها فقط مقدار غیرصفر داشته باشد.
- همیشه مقدار این تابع ≥ 0 باشد
- مجموع مقادیرشان برابر ۱ باشد.

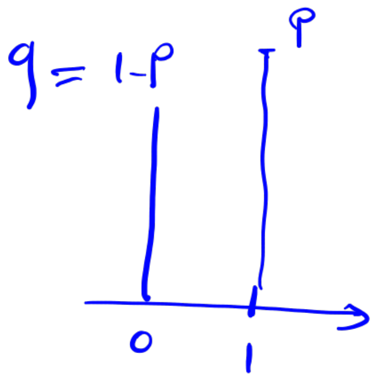
دو بعداً ✓ متغیر Binomial ✓

- برنولی را n بار تکرار کنیم و جمع آن را حساب کنیم

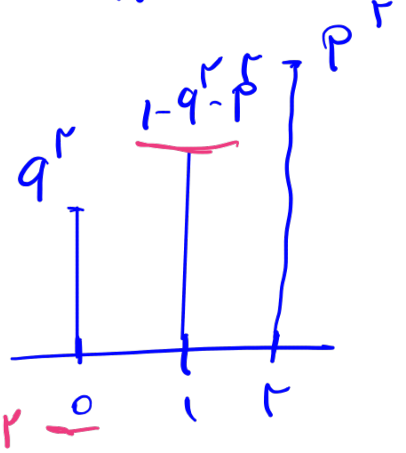
- مثلاً سکه را n بار پرتاب کنیم و بشماریم چند بار شیر آمده.

اگر که با احتمال p شیر می‌آید

$n=1$



$n=2$



$q^2 = 0$

$X \sim Binom(p, n)$

$n=2$

$P(X=\underset{k}{2}) = \binom{2}{2} p^2 q^0 = p^2$

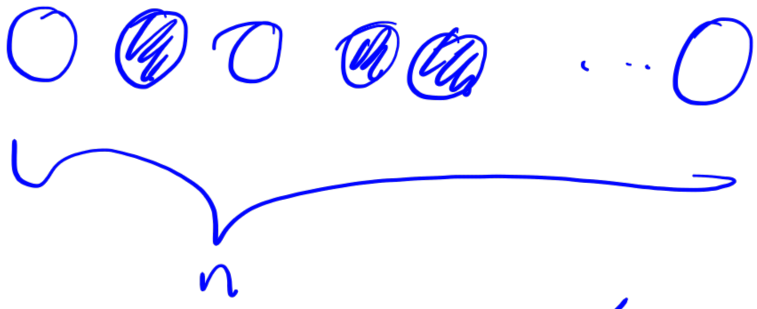
$P(X=1) = \binom{2}{1} p^1 q^1 = 2pq = 2p(1-p) = \underline{2p-2p^2}$

$1+p^2-2p$

$1-(1-p)^2-p^2 =$

$-p^2+2p-p^2 = \underline{-2p^2+2p}$

$f_X(k)=P(X=k)$ $\qquad\qquad$ $X \sim \text{Binom}(p, n)$



$\binom{n}{k} p^k q^{n-k}$

$f_X(k)$: n بار سکه‌ای که در هر بار به احتمال p جرا می‌شود را به صورت مستقل بیندازیم، چقدر احتمال دارد که دقیقاً k بار 1 بیاید.

در یک شبکه احتمال ارسال موفق هر بسته ۰٫۷ است.

اگر ۱۰ بسته ارسال کنیم چقدر احتمال دارد ۷ بسته به مقصد برسند؟

$$\binom{10}{7}\, 0.7^{7}\, 0.3^{3}$$

$$(x+y)^n = \overbrace{(x+y)(x+y)\cdots(x+y)}^{n}$$

$x =$ Success

$y =$ Failure

$$\binom{n}{k}\, x^k\, y^{n-k}$$

ضریب انتخاب

Expected Value امید ریاضی

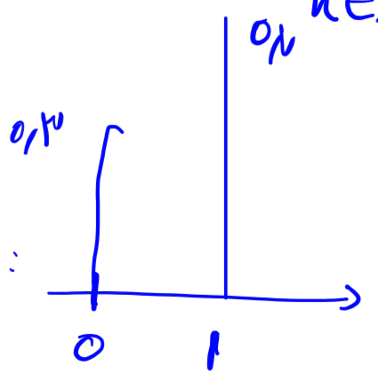
$E(X) \rightarrow$ Expected Value of X

امید ریاضی X

اگر X گسسته باشد:

$$E(X) = \sum_{x \in Support(X)} x\, P(X=x) = \sum_{x} x\, f_X(x)$$

X :



$E(X) = ۰٫۷$

می‌خواهید به یک جلسه کاری مهم برسید. اگر تاکسی اینترنتی بگیرید حتما به جلسه می‌رسید ولی اگر با اتوبوس بروید به احتمال ۰٫۳ دیر می‌رسید. دیر رسیدن به جلسه باعث خسارت ۵۰۰ هزار تومانی شما می‌شود.

هزینه اسنپ ۱۰۰ هزار تومان و اتوبوس ۵ هزار تومان است.

چه تصمیمی می‌گیرید؟

$$E(\text{Loss} \mid \text{اسنپ}) = 100000$$

$$E(\text{Loss} \mid \text{اتوبوس}) = 5000 + \underbrace{0{,}3}_{P(X=x)} \times \underbrace{500000}_{x} = 150000 + 5000 = 155000$$

در یک مسابقه تلویزیونی ۱۰ هزار نفر (با احتساب شما) شرکت می‌کنند.

Random $\left\{ \begin{array}{l} \text{یک نفر برنده ۱۰ میلیون تومان} \\ \text{۵ نفر برنده ۱ میلیون تومان می‌شوند.} \end{array} \right.$

هزینه شرکت در مسابقه چقدر پایین‌تر باشد حضور می‌یابید.

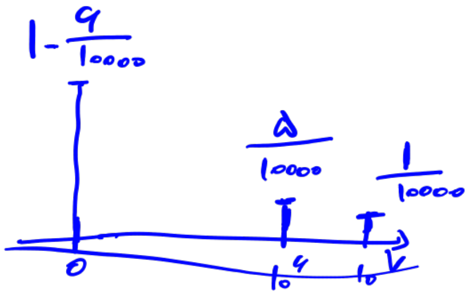
$$E(\text{gain}) = -5000 + \underbrace{E(\text{Award})}$$

$$E(\text{award}) = \sum x P(X=x) =$$

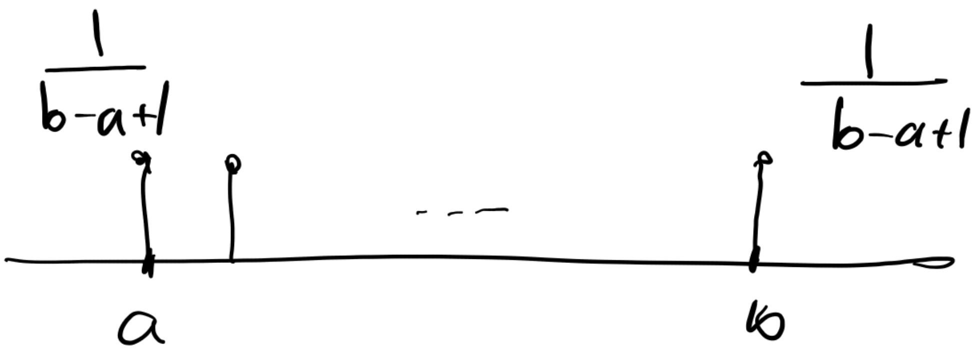
$$0 \times \cdots + \frac{10^6 \times 5}{10000} + \frac{10^7}{10000}$$

$$= 500 + 1000 = \underline{1500}$$

Award ~



Uniform یکنواخت



$a, b \in \mathbb{N}$

تاس:

$X \sim Uniform(a, b)$

$E(X) =$

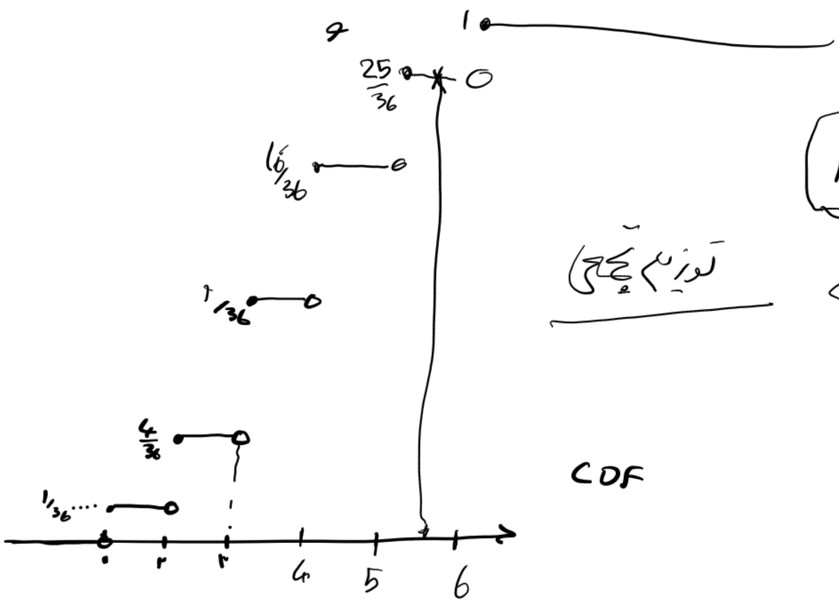
$$\frac{(b-a)^{۲}}{۱۲} \qquad \frac{a+b}{۲} \qquad \frac{1}{b-a+1}\sum_{i=a}^{b} i$$

$$\begin{array}{l} a \;\cdot\; \cdots \; b \\ \underbrace{b}_{a+b} \;\cdot\; \cdots \; a \end{array}$$

✓

$$\frac{(a+b)\cancel{(b-a+1)}}{\cancel{(b-a+1)}\;۲}$$

$\frac{1}{36}$ $\frac{4}{36}$ $\frac{9}{36}$ $\frac{16}{36}$ $\frac{25}{36}$ $\frac{36}{36}$



$\boxed{F(۵٫۵)} = \frac{۲۵}{۳۶}$

توزیع تجمعی

$P(X \leq ۵٫۵) = \frac{۲۵}{۳۶}$

CDF

$\forall x \quad F_X(x) = P(X \leq x)$

$x \to -\infty \to P(X \leq x) \to 0$

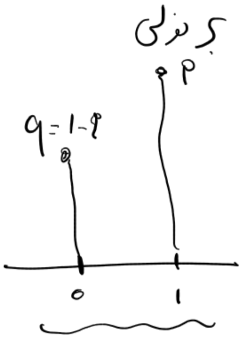
$\lim_{x \to -\infty} F(x) = 0$

$P(X \leq x)$

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

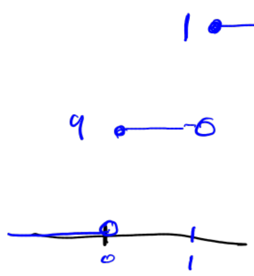
$x < y \to F(x) \leq F(y) \longrightarrow$ F صعودی است

$$P(X > x) = 1 - \underbrace{P(X \leq x)}_{F(x)} = 1 - F_X(x)$$



PMF

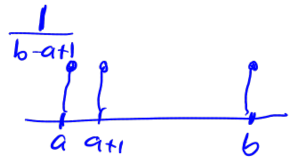
$f_X(x)$



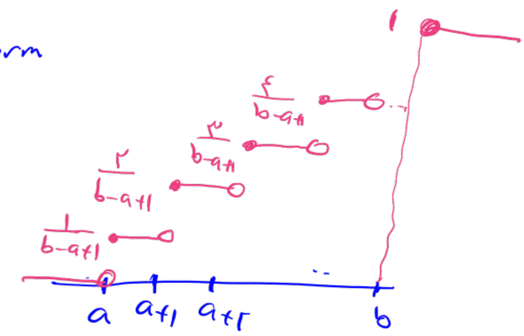
CDF

$F_X(x)$

Uniform



PMF



CDF

$$F_X(x)=\sum_{k=0}^{n}\binom{n}{k}p^k q^{n-k}$$

Binomial

دوجمله‌ای

$$F_X(a)=\int_{-\infty}^{a} f_X(x)\,dx$$

$$E(X)=\sum_{x=0}^{10} x f_X(x)=\underline{\sum x\binom{10}{x}p^x(1-p)^{10-x}}$$

خطی بودن امید ریاضی

$$E(aX) = a\,E(X)$$

a عدد ثابت

$$E(aX+b) = aE(X)+b$$

a و b ثابت   LOTUS

$$E(3X) = 3E(X) = 9$$

$$E(aX+b) = \sum_{x \in Support(X)} (ax+b)\,P(X=x) = \overbrace{a\sum x\,P(X=x) + b\underbrace{\sum_x P(X=x)}_{1}}^{aE(X)+b}$$

$$X \sim Binom(0.3,\ 10)$$

$$E(X) = \sum_{x=0}^{10} x\,P(X=x) = \sum_{x=0}^{10} x\,f_X(x) = 3$$

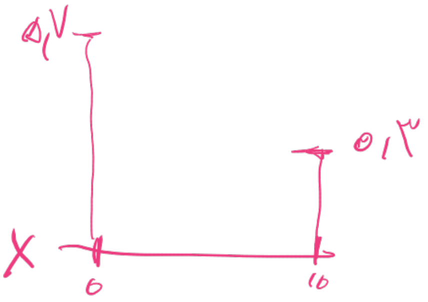
$$E(X) = 3$$

$$E(X) = E(Y+Y+\dots+Y) = \underbrace{E(Y)}_{0.3}+\underbrace{E(Y)}_{0.3}+\dots+\underbrace{E(Y)}_{0.3} = 3$$

$$X \sim Binom(0.3,10) \qquad \boxed{X = \overbrace{Y+Y+Y+\dots+Y}^{10}} \qquad Y \sim Bern(0.3)$$

$X = 10Y$ → یک سکه با احتمال ۰٫۳ را یکبار انداختم

مقدار آن را ضربدر ۱۰ کردم



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

$$E(aX+bY+c) = aE(X) + bE(Y) + c$$

a, b, c ثابت

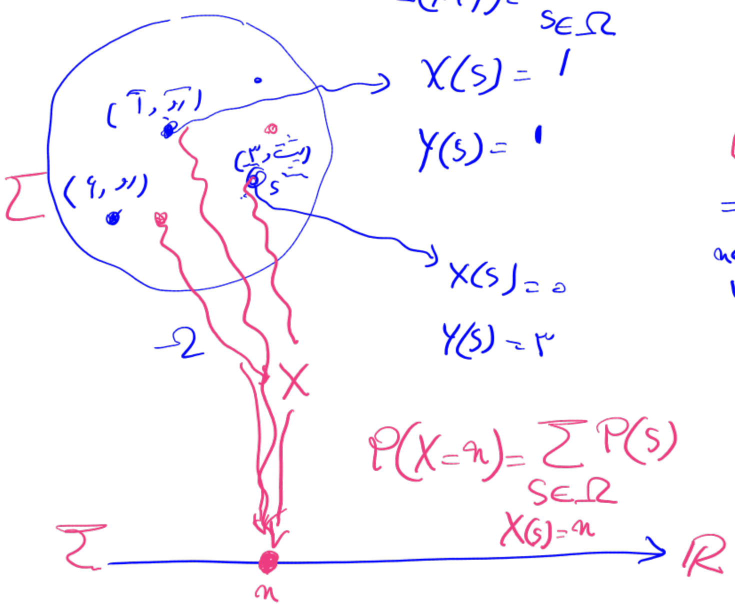
خطی بودن امید ریاضی

$$E(a_1X_1 + a_2X_2 + \cdots + a_nX_n + a_0) = a_1E(X_1) + a_2E(X_2) + \cdots + a_nE(X_n) + a_0$$

هیچ نیازی به مستقل بودن متغیرهای تصادفی نیست و در هر شرایطی برقرار است.

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

$$E(X+Y)=\sum_{s\in\Omega} P(s)(X+Y)(s)=\sum_{s} P(s)\big(X(s)+Y(s)\big)$$



$X(s)=1$

$Y(s)=1$

$X(s)=0$

$Y(s)=3$

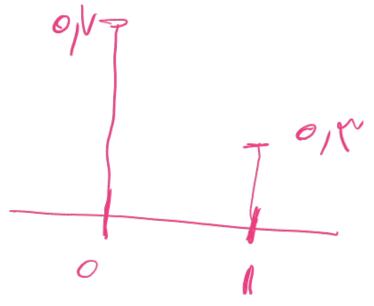
$$=\sum_{s\in\Omega} P(s)X(s)+\sum P(s)Y(s)$$

$$=\underbrace{\sum_{x\in Support(X)} x\,P(X=x)}_{E(X)}+\underbrace{\sum_{y\in Support(Y)} y\,P(Y=y)}_{E(Y)}$$

$$P(X=x)=\sum_{\substack{s\in\Omega\\ X(s)=x}} P(s)$$

$X\sim Bern(0.3)$

$$E(X^2)=\sum x^2 \overbrace{P(X=x)}^{0.3}$$



$X\sim Uniform(1,5)$ $\quad E(X^2)=?$

$$\sum_{x=1}^{5} x^2 \overbrace{P(X=x)}^{1/5}$$

$$(1^2+2^2+\dots+5^2)\frac{1}{5}$$

LOTUS

Law Of The Unconscious Statistician — تعمیم نابدیهی و نیازمند اثبات

$$E(g(X)) = \sum_{x \in \text{Support}(X)} g(x)\,P(X=x) = \sum g(x)\, f_X(x)$$

$E(Y + Z) = E(Y) + E(Z)$

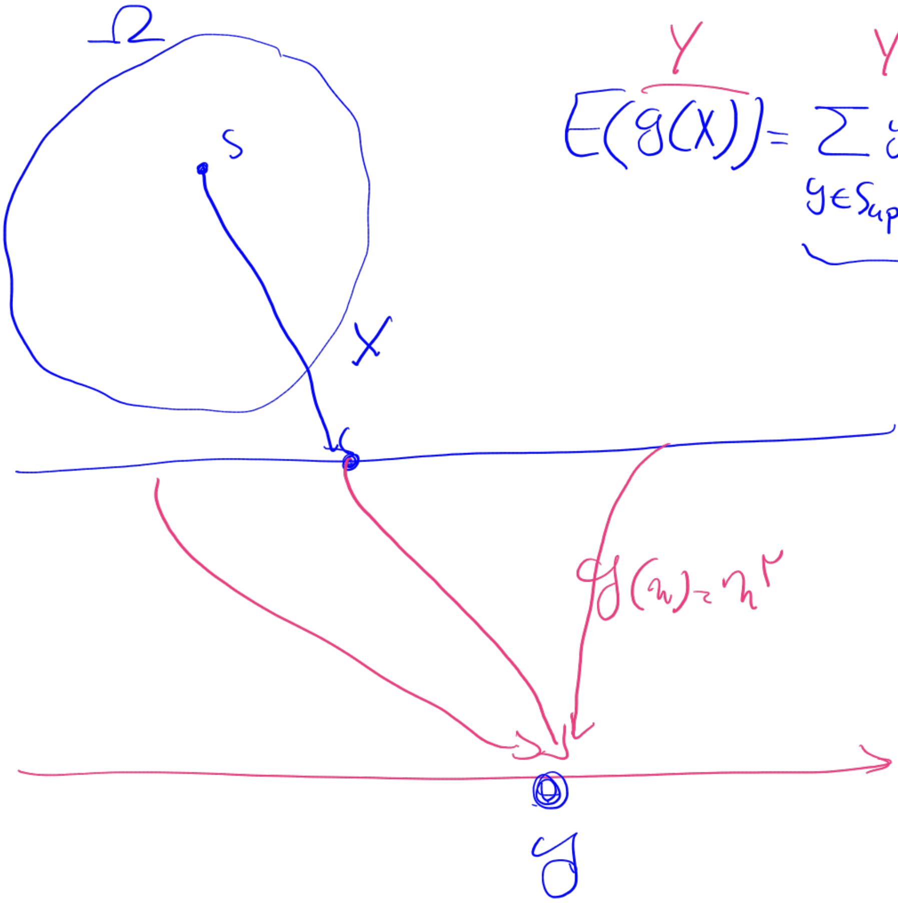
$$E(\sin(X) + X^2) = E(\sin(X)) + E(X^2)$$

خطی نیست

$\cancel{\sin(E(X))}$ $\cancel{(E(X))^2}$

$$E(X^2) = \sum_{x \in \text{Support}(X)} x^2 P(X=x)$$

$$E(\sin(X)) = \sum_{x} \sin(x)\,P(X=x)$$



$Y = g(X)$

$$E(\overbrace{g(X)}^{Y}) = \underbrace{\sum_{y \in \text{Support}(Y)} y\,P(Y=y)}_{} = \underbrace{\sum_{x \in \text{Support}(X)} g(x)\,P(Y=g(x))}_{}$$

LOTUS

Law of the Unconscious Statistician

$$E(g(X)) = \sum_x g(x)\, P(X=x) = \sum_x g(x)\, f_X(x)$$

نیاز به اثبات دارد.

X متغیر تصادفی

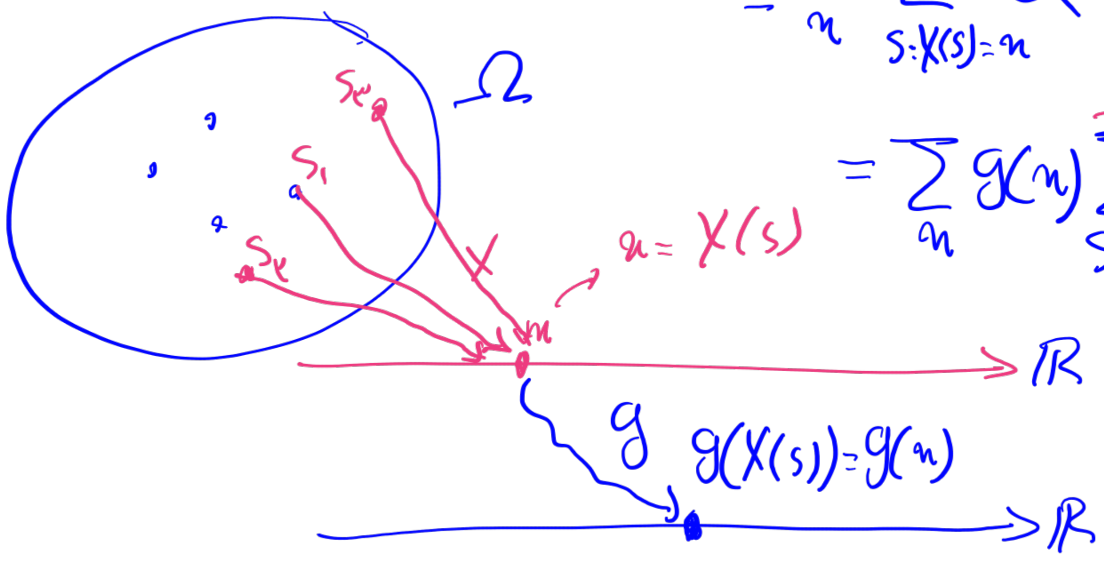
g تابع

$\mathbb{R} \to \mathbb{R}$

$$E(g(X)) = \sum_{s\in\Omega} \overbrace{g(X(s))}^{\text{شبیه یک متغیر تصادفی}} P(\{s\})$$

$$= \sum_x \sum_{s:X(s)=x} g(\overbrace{X(s)}^{x})\, P(\{s\})$$

$$= \sum_x g(x) \overbrace{\sum_{s:X(s)=x} P(\{s\})}^{P(X=x)} = \sum_x g(x) P(X=x)$$



Hypergeometric Variable

فوق هندسی

$X \sim HGeom(w, b, n)$

کیسه $\begin{cases} w \text{ تا توپ سفید} \\ b \text{ تا توپ سیاه} \end{cases}$

n توپ بدون جایگذاری بیرون می‌آوریم

• چند توپ سفید است $\leftarrow X$

$$P(X=k) = \frac{\binom{w}{k}\binom{b}{n-k}}{\binom{w+b}{n}}$$

$$E(X) = \sum_{k=0}^{\min(n,w)} k \, \frac{\binom{w}{k}\binom{b}{n-k}}{\binom{w+b}{n}}$$

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

نیازی نیست X و Y مستقل باشند

$$E(X) = E(I_1 + I_2 + \dots + I_n) = E(I_1) + E(I_2) + \dots + E(I_n) \overset{(1)}{=} \frac{nw}{w+b}$$

$$X = I_1 + I_2 + \dots + I_n$$

I_i indicator variable $\begin{cases} 1 & \text{اگر توپی که بار } i \text{ ام از کیسه بیرون می‌آوریم سفید باشد} \\ 0 & \text{در غیر این صورت (سیاه باشد)} \end{cases}$

$$E(I) = 0 \times P(I=0) + 1 \times P(I=1) = P(I=1)$$

$$E(I_2 \mid I_1 = 0) = P(I_2 = 1 \mid I_1 = 0) = \frac{w}{w+b-1}$$

$$E(I_2) = \frac{w}{w+b} \quad (1)$$

فرض کنید سال ۳۶۵ روز است. احتمال آن که در یک جامعه ۴۰ نفره حداقل دو نفر
با روز تولد یکسان حضور داشته باشند؟ (سال تولد مهم نیست)

$$1-\frac{\binom{365}{40}\times 40!}{365^{40}}$$

فرض کنید روی یک تقویم ۳۶۵ روزه هر نفر روز تولد خود را علامت می‌زند. ممکن است
یک روز چند بار علامت بخورد. تعداد متوسط روزهای علامت خورده چقدر است؟

$$E\left(I_1+I_2+\cdots+I_{365}\right)=\sum_{i=1}^{365}P(I_i=1)=365\times P(I_1=1)$$

$$I_i=\begin{cases}1 & \text{اگر روز } i \text{ ام علامت بخورد}\\ 0 & \text{نخورد}\end{cases}$$

$$=365\left(1-\left(\frac{364}{365}\right)^{40}\right)$$

$$P(I_1=1)=1-P(I_1=0)=1-\left(\frac{364}{365}\right)^{40}$$

مثل این که ۴۰ گلوله به تقویم شلیک کنیم و هیچ کدام به روز ۱ نخورد

Geometric Distribution

توزیع هندسی

$X \sim Geom(p)$ یک سکه را که با احتمال p در هر بار ۱ می‌آید آنقدر بیاندازیم که برای نخستین بار ۱ بیاید. تعداد صفرها قبل از آن $\leftarrow X$

$\underbrace{0\ 0\ 0\ 0}_{X=۴}\ 1$

تعداد تلاش ناموفق قبل از اولین تلاش موفق اگر احتمال موفقیت در هر تلاش p باشد و مستقل باشد

PMF: $P(X=k) = p(1-p)^k$

$$E(X) = \sum_{k=0}^{\infty} k\,P(X=k) = \sum_{k=0}^{\infty} k\,p(1-p)^k = p\sum_{k=0}^{\infty} k(1-p)^k = p\sum_{k=0}^{\infty} k q^k = \frac{pq}{p^2} = \frac{q}{p}$$

$p\sum_{k=0}^{\infty} k q^k =$

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

$$Q' = \frac{dQ}{dq} = 0 + 1 + 2q + 3q^2 \cdots$$

$$Q'q = q + 2q^2 + 3q^3 \cdots = \sum_{k=0}^{\infty} k q^k$$

$$q\frac{dQ}{dq} = \frac{q}{(1-q)^2} = \frac{q}{p^2}$$

$\frac{q}{1-q}$	$\frac{q^2}{1-q}$	$\frac{q^3}{1-q}$
q	q^2	q^3
q^2	q^3	
q^3		

$0<x<1 \qquad \sum_{k=0}^{\infty} x^k = 1 + x + x^2 + \cdots = \frac{1}{1-x}$

$$\sum_{i=1}^{\infty} \frac{q^i}{1-q} = \frac{1}{1-q}\underbrace{\sum_{i=1}^{\infty} q^i}_{\frac{q}{1-q}} = \frac{q}{(1-q)^2} = \frac{q}{p^2}$$

مثال: در یک شهر w نفر واکسن زدند و b نفر واکسن نزدند. اگر فرض کنیم که واکسن بر احتمال ابتلا بی‌تاثیر است و n نفر مبتلا شده باشند چقدر محتمل است که k نفر از مبتلایان واکسن زده باشند؟

$X \sim \text{HyperGeom}(w, b, n)$ ← Null Hypothesis فرض صفر

× فرض جایگزین Alternative Hypothesis ← واکسن بر احتمال ابتلا موثر است.

احتمال ابتلا را کاهش می‌دهد.

PMF(X)

$$P(X=k) = \frac{\binom{w}{k}\binom{b}{n-k}}{\binom{w+b}{n}}$$

توزیع هندسی: یک سکه با احتمال p هر بار ۱ می‌آید. چند صفر ببینیم قبل از آن که اولین بار ۱ بیاید (تعداد Failure ها قبل از اولین Success)

توزیع اولین موفقیت First Success: سکه را چند بار بیندازیم تا اولین بار ۱ بیاید. (هاش)

اگر $X \sim Geom(p)$ و تعریف کنیم $Y = 1 + X$ آنگاه $Y \sim First\ Success(p)$

$$E(Y) = 1 + E(X) = 1 + \frac{q}{p} = \frac{p+q}{p} = \frac{1}{p}$$

مثال: فرض کنید هر قوطی حاوی یک تصویر از بازیکنان فوتبال باشد و در مجموع n نفر از بازیکنان تصادفی‌شان در قوطی‌ها باشند. به طور متوسط چند قوطی را بخریم تا تصویر کل بازیکنان را داشته باشیم؟ بودن تصویر هر یک از n بازیکن در هر قوطی توزیع یکنواخت است.

تعداد کل قوطی‌هایی که باید بخریم

$$X = X_1 + X_2 + \cdots + X_n$$

هر بار که یک بازیکن جدید می‌بینیم یک موفقیت به حساب می‌آید. X_i تعداد قوطی‌هاست که بعد از $i-1$ امین موفقیت باز می‌کنیم تا موفقیت i ام به دست آید.

$X_1 = 1 \longrightarrow$ اولین قوطی حتما موفقیت است

$X_2 \sim First\ Success\left(\frac{n-1}{n}\right)$ چون تنها وقتی شکست می‌خوریم که همان بازیکن موفقیت اول را ببینیم.

$X_3 \sim First\ Success\left(\frac{n-2}{n}\right) \rightarrow E(X_3) = \frac{1}{p} = \frac{n}{n-2}$

$\vdots$

$X_n \sim First\ Success\left(\frac{1}{n}\right)$

$$E(X) = E(X_1) + E(X_2) + \cdots + E(X_n) = 1 + \frac{n}{n-1} + \frac{n}{n-2} + \cdots + n$$

توزیع هندسی: اگر سکه با احتمال p یک بیاید
چند بار صفر ببینیم قبل از اولین یک.

توزیع Negative Binomial: اگر سکه با احتمال p یک بیاید
دوجمله‌ای منفی چند بار صفر ببینیم قبل از r امین یک.

توزیع دوجمله‌ای: تعداد پرتاب ثابت بود

توزیع دوجمله‌ای منفی: تعداد موفقیت ثابت است،

مثال: تاس را چند بار بیندازیم که ۳ بار ۶ بیاید (آخرین پرتاب حساب نیست)

$X \sim \text{Negative Binom}\left(p=\frac{1}{6}, r=3\right)$

$X \sim \text{NegBinom}(p, r) \rightarrow$ PMF $f_X(k)$ $\quad P(X=k)=\binom{k}{r-1} p^r q^{k-r+1}$

r امین یک

$$\underbrace{0\ 1\ 0\ 0\ 1 \cdots 1\ 0 \cdots 0}_{k}\ 1$$

$$E(X)=\sum_{k=r-1}^{\infty} k\, P(X=k) = \sum_{k=r-1}^{\infty} k \binom{k}{r-1} p^r q^{k-r+1}$$ $\longrightarrow$ سخت است!

راه ساده‌تر محاسبه $E(X)$ وقتی $X \sim NegBinom(p, r)$:

$$X = X_1 + X_2 + \cdots + X_r$$

$X_1 \sim FirstSuccess(p)$ $\quad X_2 \sim FirstSuccess(p) \cdots \quad X_r \sim Geom(p)$

$$\underbrace{0\ 0\ 0\ 1}_{X_1}\ \underbrace{0\ 0\ \cdots\ 0\ 1}_{X_2}\ 0\ 0\ \cdots\ \underbrace{\cdots\ 1}_{X_{r-1}}\ \underbrace{0\ 0\ \cdots\ 0\ 1}_{X_r}$$

آخرین ۱ را لحاظ نمی‌کنیم.

$$E(X) = \underbrace{E(X_1)}_{\frac{1}{p}} + \underbrace{E(X_2)}_{\frac{1}{p}} + \cdots + \underbrace{E(X_{r-1})}_{\frac{1}{p}} + \underbrace{E(X_r)}_{\frac{q}{p}}$$

$$= \frac{r-1+q}{p} = \frac{q+r-1}{p}$$

متوسط تعداد پرتاب لازم قبل از این که
مثال: تاس سه بار ۶ بیاید

$X \sim NegBinom\left(\frac{1}{6}, 3\right)$ $\quad E(X) = \dfrac{\frac{5}{6} + 3 - 1}{1/6} = \dfrac{\frac{17}{6}}{\frac{1}{6}} = 17$

تعداد کل پرتاب‌ها به طور متوسط (شامل پرتاب آخر): ۱۸

توزیع پواسون — Poisson Distribution

فرض کنید رویدادی (عبور ماشین از جلوی دانشگاه، ورود یک بسته اطلاعات به یک سوئیچ شبکه، بازدید مشتری از سایت آنلاین، درخواست تاکسی اینترنتی ...) به طور متوسط λ بار در واحد زمان رخ می‌دهد. احتمال آنکه در یک واحد زمان k بار رخ دهد چقدر است؟

مثال: در هر ثانیه به طور متوسط ۲۰ درخواست تاکسی به تاکسی اینترنتی داده می‌شود. سرورهای این تاکسی حداکثر ۱۰۰ درخواست در ثانیه را پشتیبانی می‌کنند. احتمال بروز Overloading روی سرورها چقدر است؟

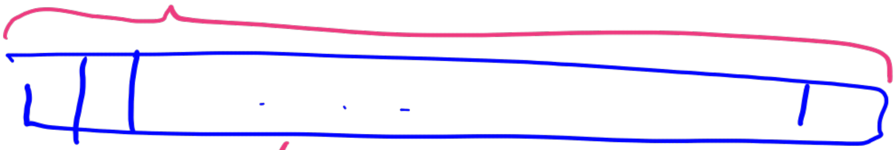
$$X \sim \text{Poisson}(\lambda = 20) \qquad P(X > 100)? = 1 - \sum_{k=0}^{100} \frac{e^{-20}\, 20^{k}}{k!}$$

$$P(X = k) = \frac{e^{-\lambda}\lambda^{k}}{k!}$$

اثبات توزیع پواسون:

Binomial

بازه زمانی به طول ۱ واحد زمان



تقسیم به n لحظه

$n \to \infty$

n بار سکه را می‌اندازیم هر بار با احتمال $P = \frac{\lambda}{n}$

$$X \sim \text{Binomial}\left(\frac{\lambda}{n}\right) \qquad \lim_{n\to\infty} P(X=k)$$

$$\lim_{n\to\infty} P(X=k) = \lim_{n\to\infty} \underbrace{\binom{n}{k}\left(\frac{\lambda}{n}\right)^k}_{A} \underbrace{\left(\frac{n-\lambda}{n}\right)^{n-k}}_{B} = \frac{\lambda^k}{k!} e^{-\lambda}$$

$$A = \lim_{n\to\infty} \frac{n(n-1)\cdots(n-k+1)}{k!} \frac{\lambda^k}{n^k} = \lim_{n\to\infty} \frac{\lambda^k}{k!} \overbrace{\frac{\overbrace{n(n-1)\cdots(n-k+1)}^{k}}{\underbrace{n \times n \times \cdots \times n}_{k}}}^{=1} = \frac{\lambda^k}{k!}$$

$$B = \lim_{n\to\infty} \left(\frac{n-\lambda}{n}\right)^{n-k} = \lim_{n\to\infty} \left(1-\frac{\lambda}{n}\right)^{\boxed{n-k}} = \lim_{n\to\infty} \left(1-\frac{\lambda}{n}\right)^{n} = e^{-\lambda}$$

$$\underbrace{\lim_{n\to\infty} \left(1-\frac{\lambda}{n}\right)^{k}}= 1$$

این جمله را می‌توانیم
ضرب در جمله بالا کنیم
تا توان n شود

بسط تیلور:

$$f(x)=f(a)+\frac{f'(a)}{1!}(x-a)+\frac{f''(a)}{2!}(x-a)^2+\frac{f'''(a)}{3!}(x-a)^3+\cdots$$

$$f(x)=e^x \quad a=0 \longrightarrow e^x=1+\frac{x}{1}+\frac{x^2}{2!}+\frac{x^3}{3!}+\cdots=\sum_{k=0}^{\infty}\frac{x^k}{k!}$$

$$\left(1-\frac{\lambda}{n}\right)^n=\sum \underbrace{\binom{n}{k}1^{n-k}}\left(\frac{-\lambda}{n}\right)^k$$

$$=\sum \frac{\cancel{n(n-1)\cdots(n-k+1)}}{k!}\frac{(-\lambda)^k}{\cancel{n^k}}=\underbrace{e^{-\lambda}}$$

$n\to\infty$

Variance

Moment

$$\underbrace{\sum_x x^i P(X=x)}_{\text{raw Moment امین } i}$$

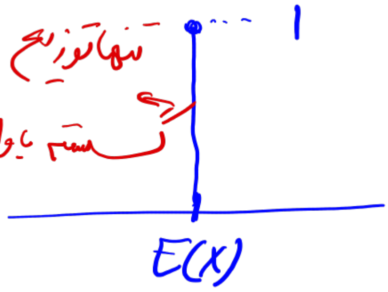
$$\underbrace{\sum_x (x-E(X))^i P(X=x)}_{\text{Central Moment امین } i}$$

Moment شماره	raw	Central
۱	Expected Value	0
۲		Variance

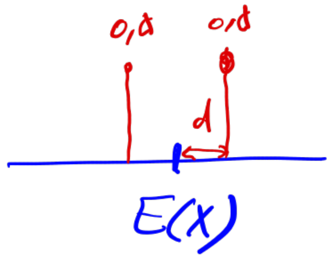
$$\underbrace{\sum_x (x-E(X))P(X=x)}_{\text{Central Momentum اولین}} = \underbrace{\sum_x x P(X=x)}_{E(X)} - \underbrace{\sum \underbrace{E(X)}_{\text{عدد ثابت}} P(X=x)}_{E(X)\times 1} = 0$$

$$Var(X) = \underbrace{\sum (x-E(X))^2 P(X=x)}_{\text{Central Momentum دومین}}$$

تنها توزیع مستقیم با واریانس صفر



$Var(X) = ۰٫۵ \times d^2 + ۰٫۵ \times d^2 = d^2$



$X \sim Bern(P)$

$$Var(X) = \sum_x (x-\underbrace{E(X)}_{P})^2 P(X=x) = P^2 \times (1-P) + (1-P)^2 \times P =$$

$$P(1-P)(\underbrace{P+1-P}_{1}) = Pq$$

$$\begin{cases} P=1 \rightarrow Var(X)=0 \\ q=1 \ (P=0) \rightarrow Var(X)=0 \\ \underbrace{\arg\max_P (Var(X))}_{\text{p را چه مقداری بدهیم تا Var(X) بیشترین شود}} = \frac{1}{۲} \rightarrow P=1 \Rightarrow Var(X) = \frac{1}{۴} \end{cases}$$

$X \sim \text{Binom}(p, n) \qquad E(X) = np$

$$Var(X) = \sum_{k=0}^{n} \Big[\underbrace{\binom{n}{k} p^{k} (1-p)^{n-k}}_{P(X=k)} \Big] \underbrace{(k - np)^2}_{(x - E(X))^2}$$

تفسیر:

$$Var(X) = E\Big(\underbrace{(X - E(X))^2}_{\text{متغیر تصادفی جدید}}\Big) \quad : \quad g(X) = (X - E(X))^2$$

$$\underset{\text{LOTUS}}{E[g(X)]} = E\big((X - E(X))^2\big) \underset{\text{Lotus}}{=} \sum_{x} g(x) P(X=x) = \underbrace{\sum \big((x - E(X))^2\big) P(X=x)}_{Var(X)}$$

$$E\big((X - E(X))^2\big) = E\Big[X^2 + \underbrace{E(X)^2}_{\text{عدد ثابت}} - 2 X E(X)\Big] = E(X^2) + E(X)^2 - 2E\big[X \underbrace{E(X)}_{\text{ثابت}}\big]$$

تفسیر:

خطی بودن امید ریاضی

$$= E(X^2) + E(X)^2 - 2E(X)^2$$

$$Var(X) = E(X^2) - E(X)^2$$

$E(X^2) \rightarrow$ از متغیر تصادفی X^2 امید ریاضی بگیریم $= \sum x^2 P(X=x)$ LOTUS

$E(X)^2 \longrightarrow$ $E(X)$ را حساب و بعد به توان ۲ برسانیم

توضیح:

$$E(X \underbrace{E(X)}_{\text{عدد ثابت}}) = E(X)E(X) = E(X)^2$$

توضیح

$$Var(X+Y) = \underbrace{E((X+Y)^2)}_{\underbrace{E(X^2+Y^2+2XY)}_{E(X^2)+E(Y^2)+2E(XY)}} - \underbrace{E(X+Y)^2}_{\underbrace{[E(X)+E(Y)]^2}_{E(X)^2+E(Y)^2+2E(X)E(Y)}} =$$

حاصل
$$= \underbrace{[E(X^2)-E(X)^2]}_{Var(X)} + \underbrace{[E(Y^2)-E(Y)^2]}_{Var(Y)} + 2\underbrace{[E(XY)-E(X)E(Y)]}_{Cov(X,Y)}$$

$$Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$$

اگر X, Y مستقل باشند:

$$Cov(X,Y) = E(XY) - E(X)E(Y)$$

LOTUS: $$\sum_{x,y} xy \underbrace{P(X=x, Y=y)}_{\text{مستقل بودن } X,Y} = \sum_{x,y} xy\, P(X=x)P(Y=y) = \underbrace{\sum_x xP(X=x)}_{E(X)} \underbrace{\sum_y yP(Y=y)}_{E(Y)}$$

X, Y مستقل

$$\Rightarrow E(XY) = E(X)E(Y) \rightarrow Cov(X,Y) = 0$$

$$\Rightarrow Var(X+Y) = Var(X) + Var(Y)$$

مثال:

$Var(X)$ $\qquad X \sim Binom(p, n)$

$$X = X_1 + X_2 + \cdots + X_n$$

X_i ها independent and identically Distributed

$$\underbrace{i.i.d}$$

$$\underbrace{X_i \sim Bern(p)}_{\text{identically distributed}} \qquad \underbrace{X_i \text{ ها مستقلند}}_{\text{independent}}$$

چون X_i ها مستقلند :

$$Var(X) = \underbrace{Var(X_1)}_{pq} + \underbrace{Var(X_2)}_{pq} + \cdots + \underbrace{Var(X_n)}_{pq}$$

$$= npq$$

$X \sim Uniform(a, b)$ $\qquad E(X) = \frac{a+b}{2}$

$$Var(X) = \sum_{x=a}^{b} \left(x - \frac{a+b}{2}\right)^2 \frac{1}{b-a+1}$$

= تمرین برای خودتان

$X \sim \text{Poisson}(\lambda)$

$$E(X) = \sum_{k=0}^{\infty} k P(X=k) = \sum k \underbrace{\frac{e^{-\lambda}\lambda^k}{k!}}$$

بسط تیلور

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)(x-a)^2}{2!} + \frac{f'''(a)(x-a)^3}{3!} + \cdots = \sum_{k=0}^{\infty} \frac{f^{(k)}(a)(x-a)^k}{k!}$$

$a=0$

$$\underbrace{e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}}$$

براساس x مشتق بگیریم

$$\rightarrow \sum \frac{e^{-\lambda}\lambda^k}{k!} = e^{-\lambda} \underbrace{\sum_{k=0}^{\infty} \frac{\lambda^k}{k!}}_{e^{\lambda}} = 1$$

حاصل جمع PMF توزیع پواسون برابر یک

$$\underbrace{e^x = \sum_{k=0}^{\infty} k \frac{x^{k-1}}{k!}}$$

اگر $x=\lambda$ و دو طرف در $e^{-\lambda}\lambda$ ضرب کنیم

$$\rightarrow \underbrace{e^{-\lambda}e^{\lambda}}_{1}\lambda = \underbrace{\sum_{k=0}^{\infty} k \frac{e^{-\lambda}\lambda^k}{k!}}_{E(X)} = \lambda$$

$$\Rightarrow E(X) = \lambda$$

$$Var(X) = \sum_{k=0}^{\infty} (k-\lambda)^2 P(X=k) = \sum (k-\lambda)^2 \underbrace{\frac{e^{-\lambda}\lambda^k}{k!}}$$

محاسبه دشوار

روش دوم:

$$Var(X) = E(X^2) - \underbrace{E(X)^2}_{\lambda^2}$$

$$E(X^2)=\sum_{k=0}^{\infty} k^2 P(X=k) = \sum_{k=0}^{\infty} k^{\cancel{2}} \frac{e^{-\lambda}\lambda^k}{\cancel{k!}\;(k-1)!}$$

LOTUS

$$= e^{-\lambda}\sum_{k=1}^{\infty} \underset{k-1+1}{k}\,\frac{\lambda^k}{(k-1)!}$$

$$=\lambda e^{-\lambda}\left[\underbrace{\lambda\sum_{k=1}^{\infty} (k-1)\frac{\lambda^{k-2}}{(k-1)!}}_{k'=k-1} + \underbrace{\sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!}}_{k'=k-1 \;\to\; \underbrace{\sum_{k'=0}^{\infty}\frac{\lambda^{k'}}{k'!}}_{e^{\lambda}}}\right]$$

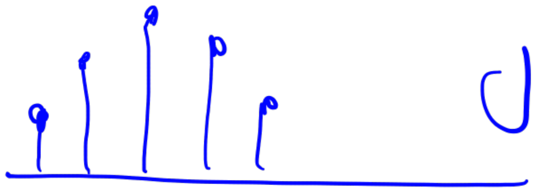
کل این عبارت e^{λ} است

$$\lambda e^{-\lambda}\left[\lambda e^{\lambda} + e^{\lambda}\right] = \lambda \underbrace{e^{-\lambda} e^{\lambda}}_{1}\left[\lambda+1\right] = \lambda^2+\lambda$$

$$E(X^2)-E(X)^2 = \lambda^2+\lambda-\lambda^2 = \underline{\lambda}$$

$$X\sim \text{Poisson}(\lambda)\begin{cases} E(X)=\lambda \\ Var(X)=\lambda \end{cases}$$

متغیرهای تصادفی پیوسته:



جرم احتمال

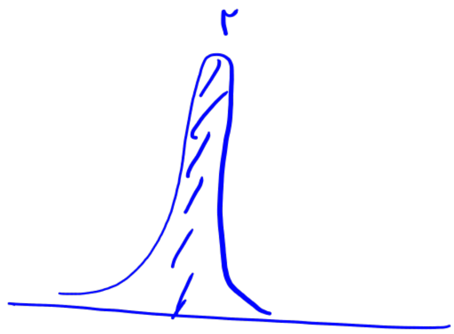
PMF

تابع چگالی احتمال

احتمال قد

Probability Density Function

$$f_X(x)$$



$$P(a \leq X \leq b) = \int_a^b f_X(x)\,dx$$

خصوصیات $f_X(x)$:

۱- $\forall_{x \in R} \quad f_X(x) \geq 0$

۲- $\int_{-\infty}^{\infty} f_X(x) = 1$

به ازای هر a و b ها، احتمال پیشامد آنکه $a \leq X \leq b$ باشد برابر $\int_a^b f_X(x)$

تعریف تابع چگالی احتمال

برای متغیر تصادفی پیوسته :

احتمال $X=a$

$$P(X=a) = 0$$

a عدد ثابت

مثال: احتمال اینکه قد یک نفر دقیقاً 140 cm باشد یعنی

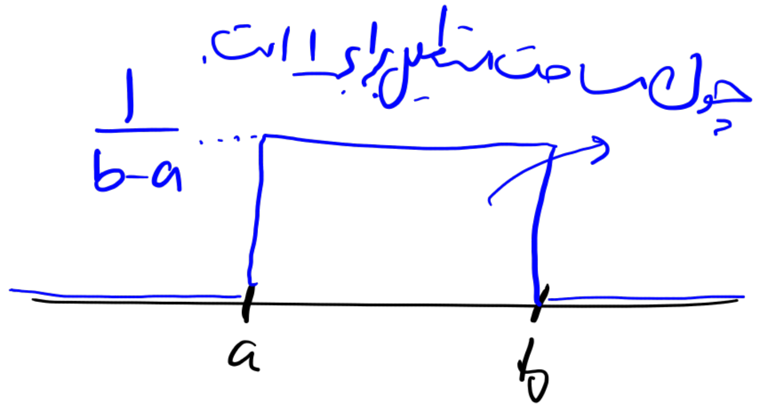
$140,00000$ ...

همیشه احتمال متغیر تصادفی پیوسته را روی یک بازه یا مجموعه (بی‌نهایت عضو) تعریف می‌کنیم

مثال: توزیع یکنواخت از a تا b:

$$X \sim \text{Uniform}(a,b)$$

$$f_X(x) =$$



$$E(X) = \int_{-\infty}^{+\infty} x f_X(x)\,dx \Rightarrow$$

$$E(X)_{X \sim \text{Uniform}(a,b)} = \int_a^b x \frac{1}{b-a}\,dx = \frac{x^2}{2(b-a)}\Big|_a^b = \frac{b^2-a^2}{2(b-a)}$$

$$= \frac{a+b}{2}$$

یادآوری: توزیع یکنواخت پیوسته

$$X \sim \overset{\text{پیوسته}}{\text{Unif}}(a,b)$$

$$E(X) = \int_{-\infty}^{\infty} x f_X(x)\,dx = \int_a^b \frac{1}{b-a} x\,dx = \frac{x^2}{2(b-a)}\Big|_a^b = \frac{b^2-a^2}{2(b-a)} = \frac{a+b}{2}$$

$$f_X(x) = \begin{cases} \frac{1}{b-a} & a \le x \le b \\ 0 & \text{o.w.} \end{cases}$$

$$Var(X) = \int_{-\infty}^{+\infty} (x - E(X))^2 f_X(x)\,dx = \int_a^b \left(x - \frac{a+b}{2}\right)^2 \frac{1}{b-a}\,dx$$

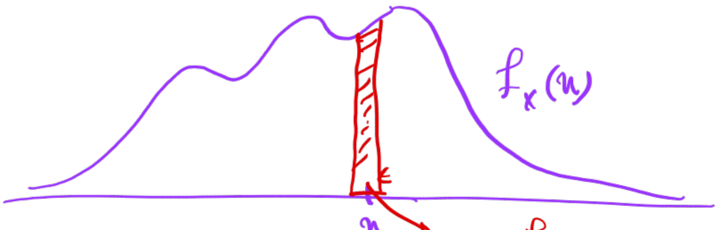
$$Var(X) = E(X^2) - \underbrace{E(X)^2}_{\left(\frac{b+a}{2}\right)^2}$$

$$E(X^2) \underset{\text{LOTUS}}{=} \int_a^b x^2 \frac{1}{b-a}\,dx = \frac{x^3}{3(b-a)}\Big|_a^b = \frac{b^3-a^3}{3(b-a)} = \frac{a^2+ab+b^2}{3}$$

$$E(g(X)) \underset{\text{LOTUS}}{=} \int_{-\infty}^{+\infty} g(x) f_X(x)\,dx$$

LOTUS برای متغیر تصادفی پیوسته

$$Var(X) = \frac{a^2+ab+b^2}{3} - \frac{a^2+2ab+b^2}{4} = \frac{a^2-2ab+b^2}{12} = \frac{(b-a)^2}{12}$$



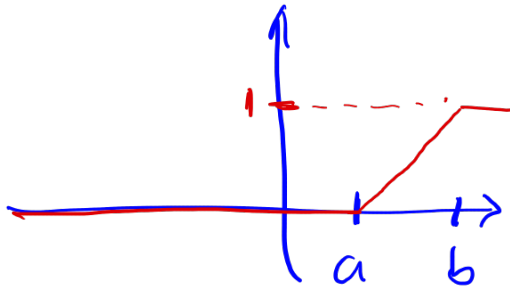
$P(x - \frac{\epsilon}{2} \le X \le x + \frac{\epsilon}{2}) =$ مساحت $\sim \epsilon f_X(x)$

$$F_X(x) = P(X \le x) = \int_{-\infty}^{x} f_X(u)\,du = \int_a^x \frac{dx}{b-a}$$

$x < a \rightarrow F_X(x) = 0$

$a \le x \le b$

$$F_X(x) = \frac{u}{b-a}\Big|_a^x = \frac{x-a}{b-a}$$

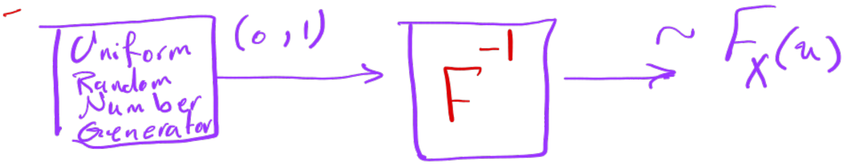


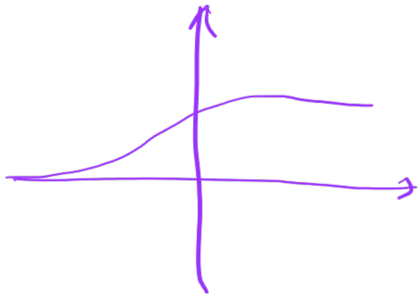
$$x > b \rightarrow F_X(x) = 1$$

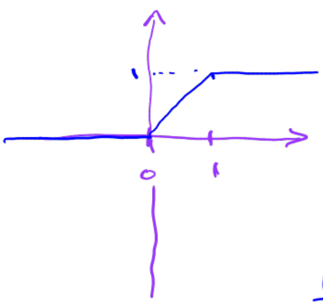
$$X \sim Unif(a,b) \rightarrow F_X(x) \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a < x < b \\ 1 & b < x \end{cases}$$

Universality of Uniform Distribution

سوال: فرض کنید من یک تولیدکننده اعداد تصادفی دارم که از Uniform(0,1) نمونه تولید می‌کند.
حالا چطور با استفاده از آن، اعداد تصادفی از یک توزیع دلخواه شبیه‌سازی کنم؟







فرض کنیم F اکیدا صعودی است

CDF

$F(x)$

$F^{-1}(x)$

$F^{-1}(x)$

Vorası

$$P(U \leq a) = a$$

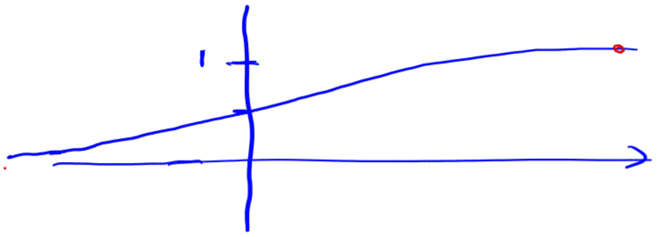
$U \sim Uniform(0,1)$ $\quad \forall x \in \mathbb{R}$ $\quad$ چون F صعودی است $\quad 0 \leq F(x) \leq 1$

$X = F^{-1}(U)$ $\quad U \leq F(x)$

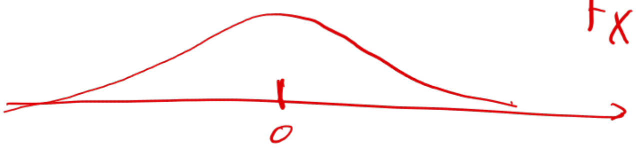
$$P(X \leq x) = P(F^{-1}(U) \leq x) = P(F(F^{-1}(U)) \leq F(x)) = F(x)$$

$X \sim F$

مثال: $F = \dfrac{e^x}{1+e^x}$ سیگموید



CDF



$f_X(x)$ PDF

$$y = \frac{e^x}{1+e^x} \rightarrow y + ye^x = e^x \qquad e^x(1-y) = y \rightarrow e^x = \frac{y}{1-y}$$

$$x = \ln\left(\frac{y}{1-y}\right)$$

$$y = F(x)$$

$$x = F^{-1}(y) = \ln\left(\frac{y}{1-y}\right)$$

$$U \sim \text{Uniform}(0,1) \rightarrow \ln\left(\frac{U}{1-U}\right) \sim \text{Sigmoid}$$

توزیع نمایی Exponential

اگر به طور متوسط در هر واحد زمان λ بار یک اتفاق بیافتد، چقدر باید صبر کنیم تا اولین بار آن اتفاق رخ دهد؟

مثال: در ایران سالانه حدود یک میلیون نوزاد متولد می‌شود. از همین لحظه چقدر باید صبر کنیم تا نخستین نوزاد بعدی متولد شود.

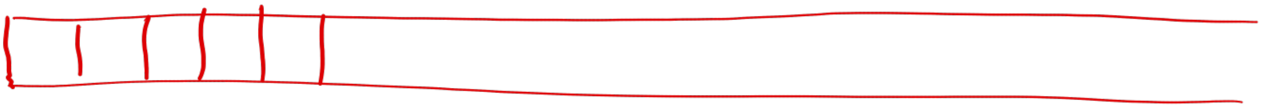
زمان ← دارد از یک توزیع تبعیت می‌کند.

← احتمال آنکه زمان انتظار تا تولد نخستین نوزاد بعدی بیش از یک ساعت باشد چقدر است؟

$\text{Bern} \xrightarrow{\text{تکرار}} \text{Binomial} \xrightarrow{\lim_{n\to\infty}} \text{Poisson}$

به جای شمردن تعداد ۱ها تعداد دفعات تا اولین ۱ را شمردیم

$\text{Geometric} \xrightarrow[\lim_{n\to\infty}]{} \text{Exponential}$



در هر واحد زمان n بار سکه می‌اندازیم. با احتمال $p=\frac{\lambda}{n}$ سکه یک می‌آید. X را متغیر تصادفی تعداد انداختن سکه تا قبل از آمدن اولین یک در نظر می‌گیریم. اگر $n\to\infty$، X از چه توزیعی پیروی می‌کند؟

λ: متوسط تعداد وقوع اتفاق در هر واحد زمان

واحد زمان: سال $\qquad \lambda = 1000000$

$$P(X>\underset{\text{زمان}}{t}) = (1-p)^{tn}$$ → در tn بار اول پرتاب سکه همیشه صفر باید بیاید

$X \sim \text{Geometric}$

تعداد دفعات سکه انداختن در x زمان: $x\,n$

$$\lim_{n\to\infty}(1-p)^{tn} = \underbrace{\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{tn}}_{} \rightarrow \lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{\frac{n}{\lambda}\lambda t} = \left(\frac{1}{e}\right)^{\lambda t} = e^{-\lambda t} \quad (1)$$

$$w=\frac{n}{\lambda} \rightarrow \frac{1}{w}=\frac{\lambda}{n}$$

$$\lim_{w\to\infty}\left(1-\frac{1}{w}\right)^{w} = \frac{1}{e} \qquad \lim_{w\to\infty}\left(1+\frac{1}{w}\right)^{w} = e$$

$$(1) \rightarrow P(X\le t) = \underbrace{1-e^{-\lambda t}}_{\text{CDF}} \Rightarrow F_X(t)$$

$t<0 \rightarrow F_X(t)=0$

تابع توزیع

$$\underbrace{f_X(t) = \lambda e^{-\lambda t}}_{\text{PDF}}$$

توزیع نمایی

فرض:

مثال: تولید مولد عدد تصادفی که از توزیع نمایی تبعیت کند:

$$t > 0 \rightarrow y = F_X(t) = 1 - e^{-\lambda t} \rightarrow 1 - y = e^{-\lambda t}$$

$$\ln(1-y) = -\lambda t$$

$$F^{-1}(y) = t = \frac{-\ln(1-y)}{\lambda}$$

CDF

$$P_{(X \le x)} = \overbrace{F_X(x)} = \int_{-\infty}^{x} \underbrace{f_X(t)}_{\text{PDF}}\, dt$$

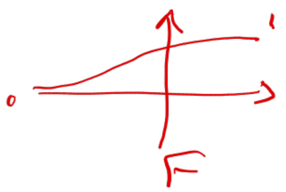
برای همه توزیع های پیوسته:

Commulative Distribution function

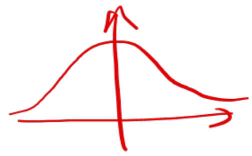
F انتگرال f است

f مشتق F است.

پیوسته

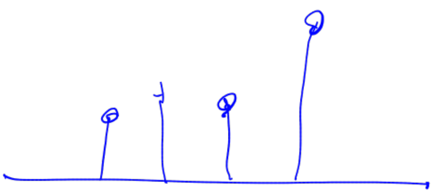


F



f

Probability Density Function



F : CDF

گسسته:

f : Probability Mass Function

$X \sim \text{Exponential}(\lambda)$

Exponential

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

$$E(X) = \int_{-\infty}^{+\infty} x f_X(x)\, dx$$

$$= \underbrace{\int_0^{\infty} \underbrace{x}_{g} \underbrace{\lambda e^{-\lambda x}}_{f'} dx}_{\text{جزء به جزء}} = -x e^{-\lambda x} + \underbrace{\int e^{-\lambda x} dx}_{-\frac{1}{\lambda} e^{-\lambda x}} = \underbrace{-e^{-\lambda x}\left[x + \frac{1}{\lambda}\right]\Big|_0^{\infty}}$$

$$f = -e^{-\lambda x}$$

$$\int u\, dv = uv - \int v\, du$$

$$\int f' g = fg - \int g' f$$

$$\underbrace{-x e^{-\lambda x}\Big|_0^{\infty}}_{0} + \frac{e^{-\lambda x}}{\lambda}\Big|_0^{\infty} = \frac{1}{\lambda}$$

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

$$Var(X) = E(X^2) - E(X)^2 \Rightarrow ?$$

$$E(X^2) \underset{\text{LOTUS}}{=} \int_{-\infty}^{+\infty} x^2 f_X(x)\, dx = \int_0^{+\infty} x^2 \lambda e^{-\lambda x} dx \underset{\text{جزء به جزء}}{=} \frac{2}{\lambda^2}$$

$$\frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2} = Var(X)$$

عدم حافظه بودن توزیع نمایی:

← هندسی: اگر بگویند تا الان شکم ۱۰ بار صفر آمده از الان چند بار باید بیندازیم تا شیر بیاید؟

A

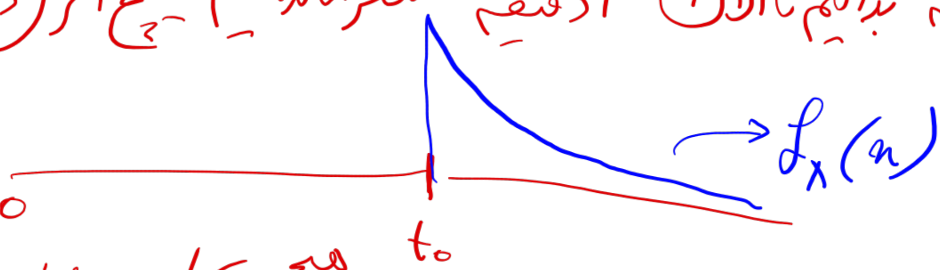
B

B مستقل از A است ← توزیع هندسی بی حافظه است

در مورد توزیع نمایی هم عین همین است:

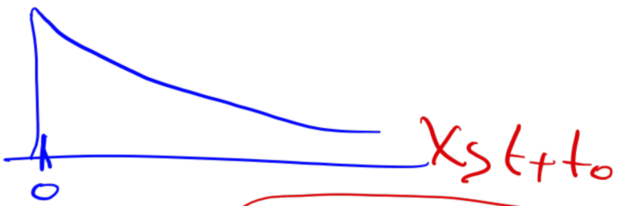
مثلاً اگر فرض کنیم زمان قبول کردن تاکسی اینترنتی از توزیع نمایی پیروی می‌کند

این که بدانیم تا الان ۳ دقیقه منتظر مانده‌ایم هیچ اثری در زمان قبول ندارد



هیچ بجای ۳ دقیقه نیامد

$t_0 > 0$



$$P(X > t+t_0 \mid X > t_0) = \frac{P(X > t+t_0 \cap X > t_0)}{P(X > t_0)} = \frac{P(X > t+t_0)}{P(X > t_0)}$$

$$= \frac{e^{-\lambda(t+t_0)}}{e^{-\lambda t_0}} = e^{-\lambda t} = P(X > t)$$

یعنی اگر مبدأ زمان را در همین لحظه بگذاریم و هیچ اطلاعی از این که در زمان t_0 قبلی ما اتفاقی نیافتاده است.

فرض کنید $X_1 \dots X_n \sim D$ و $X_1 \dots X_n$ مستقل

$$\underbrace{X_1 \dots X_n \sim D}$$

$E(D) = E(X_i) = \mu$

$Var(X_i) = \sigma^2$

i i d

independent, identically Distributed

$$\boxed{Y = X_1 + X_2 + \dots X_n}$$

$E(Y) = E(X_1 + \dots + X_n) = E(X_1) + E(X_2) \dots + E(X_n) = n\mu$

n زیاد

$n \to \infty$

$Var(Y) = Var(X_1) + Var(X_2) + \dots + Var(X_n)$

چون X_i ها مستقل هستند

$Var(Y) = n\sigma^2$

Y

از چه توزیعی پیروی می‌کند؟

$$Z = \frac{Y - n\mu}{\sqrt{n}\,\sigma}$$

$$E(Z) = E\left(\frac{Y - n\mu}{\sqrt{n}\,\sigma}\right) = \frac{1}{\sqrt{n}\,\sigma} E(Y - n\mu) = 0$$

هر عدد ثابت دلخواه

هر متغیر تصادفی دلخواه

$$Var(aX) = E\left(\left(aX - E(aX)\right)^2\right) =$$

$$\underbrace{E\left((aX)^2\right)}_{a^2X^2} + \underbrace{E(aX)^2}_{(aE(X))^2} - 2aX\underbrace{E(aX)}_{aE(X)} =$$

$$a^2E(X^2) - a^2E(X)^2 = a^2Var(X)$$

هر متغیر تصادفی دلخواه

هر عدد ثابت

$$Var(X + a) = Var(X)$$

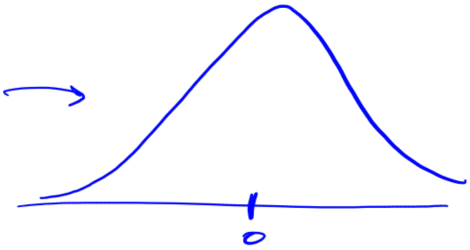
$$Var(aX) = a^2 Var(X)$$

$$Var(Z) = Var\left(\frac{Y - n\mu}{\underbrace{\sqrt{n}\,\sigma}_{\text{ضریب ثابت}}}\right) = \frac{1}{n\sigma^2} Var(Y - \overbrace{n\mu}^{\text{عدد ثابت}}) = \frac{1}{n\sigma^2} \overbrace{Var(Y)}^{n\sigma^2} = 1$$

(کلاً به ازای هر متغیر تصادفی دلخواه X با هر توزیع دلخواه مبحث می‌کند اگر $Z = \frac{X - E(X)}{\sqrt{Var(X)}}$ آن‌گاه $E(Z) = 0$, $Var(Z) = 1$)

$Y = X_1 + X_2 + \cdots X_n$

$$\frac{Y - E(Y)}{\sqrt{Var(Y)}} = Z \to$$



نرمال استاندارد

$n \to \infty$

$E(Z) = 0 \qquad Var(Z) = 1$

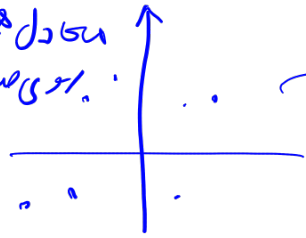
$$f_Z(z) \sim \boxed{e^{-\frac{z^2}{2}}}$$

$$\int_{-\infty}^{+\infty} e^{-\frac{z^2}{2}} = \sqrt{2\pi}$$

$$\varphi(z) = f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \to \int_{-\infty}^{+\infty} f_Z(z)\,dz = 1$$

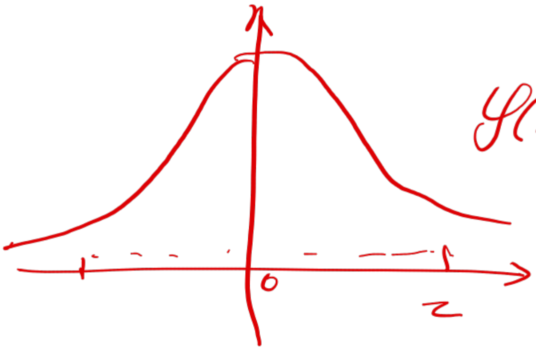
$$\int_{-\infty}^{+\infty} e^{-\frac{z^2}{2}} dz \int_{-\infty}^{+\infty} e^{-\frac{w^2}{2}} dw = \iint_{-\infty}^{+\infty} e^{-\frac{z^2}{2}} e^{-\frac{w^2}{2}} dw\,dz$$

معادل همه نقاط
روی صفحه دوبعدی



Polar $\to = 2\pi$

r, θ



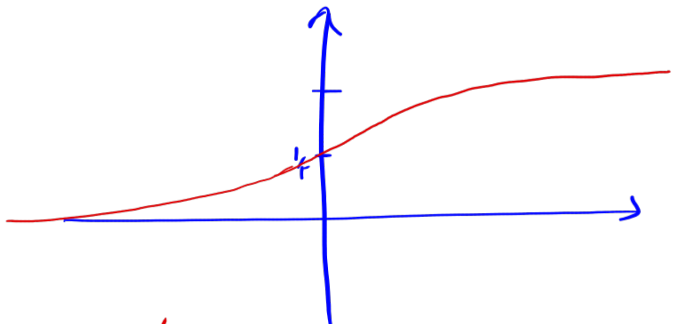
$$\varphi(z) = \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}}$$

متقارن
حول صفر

$\varphi(z) = \varphi(-z) \to$ PDF تابع چگالی احتمال

$$\underline{\Phi(z) = F_Z(z)} = \int_{-\infty}^{z} \varphi(x)\,dx = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

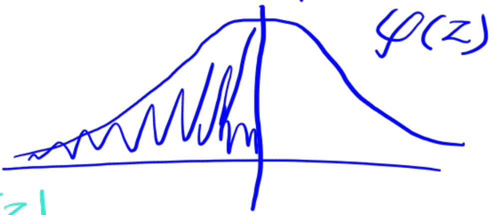
CDF
تابع احتمال تجمعی



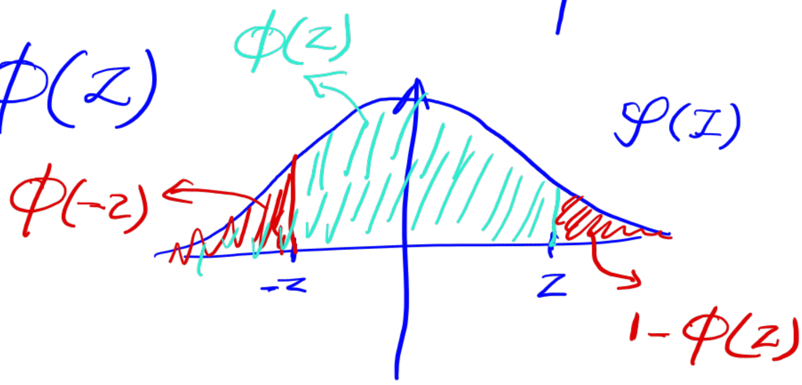
$\Phi(Z)$

شکل بسته از تابع $\Phi(Z)$ موجود نیست.

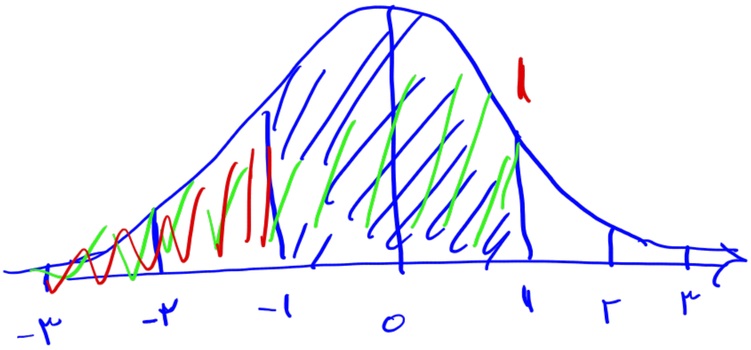
$$\lim_{x\to-\infty}\Phi(x)=0 \qquad \lim_{x\to+\infty}\Phi(x)=1 \qquad \Phi(0)=1/2$$



$$\underbrace{\Phi(-z)}_{}=1-\Phi(z)$$



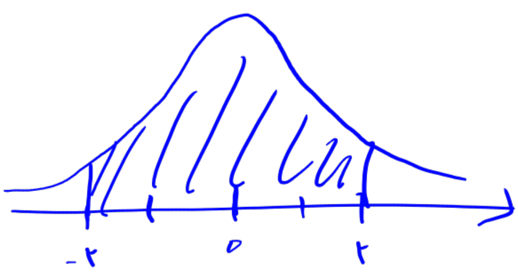
مثال :



$$\int_{-1}^{1}\varphi(z)\,dz=\int_{-1}^{1}\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}dz=\Phi(1)-\overbrace{\underbrace{\Phi(-1)}_{}}^{1-\Phi(1)}=2\Phi(1)-1=\boxed{0.68}$$

$$\Phi(1)=0.8413\to 2\Phi(1)-1=\boxed{0.68}$$

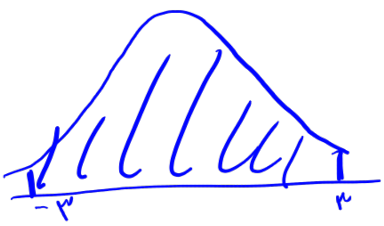
$$1.6826-1=$$



$$\int_{-2}^{2}\varphi(z)\,dz=2\Phi(2)-1=\boxed{0.95}$$

$$\Phi(2)=0.9773$$

$$2\Phi(2)-1=0.9546$$



$$\int_{-3}^{3}\varphi(z)\,dz=2\Phi(3)-1=0.997$$

قاعده ۶۸ - ۹۵ - ۹۹,۷

هرگاه یک متغیر نرمال استاندارد Z داشته باشیم

$Z \sim N(0,1)$

$\sigma = 1 \to$ انحراف معیار

$\sigma^2 = 1 \to$ واریانس

$$P(-1 \leq Z \leq 1) = P(-\sigma \leq Z \leq \sigma) = 68\%$$

$$P(-2 \leq Z \leq 2) = 95\%$$

$$P(-3 \leq Z \leq 3) = 99.7\%$$

توزیع نرمال در حالت کلی:

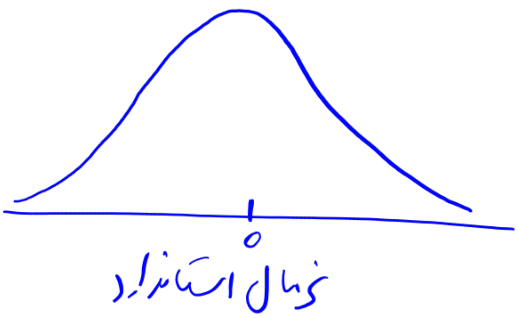
$X \sim N(\mu, \sigma^2)$

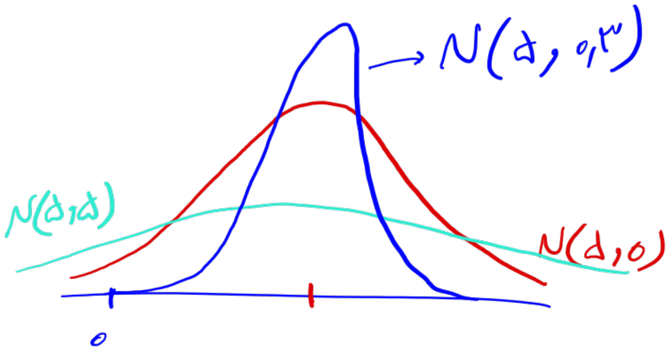
$$E(X) = \mu$$

$$Var(X) = \sigma^2$$

PDF: $$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$F_X(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt \to$$ فرم بسته ندارد.





Moments

$$E(X^n) = M_n(X) \qquad M_1(X) = E(X) = \text{Expected Value}$$

$$\underbrace{E(X^n)}_{\text{Raw Moment}}$$

$$\underbrace{E\left(\left(X-E(X)\right)^n\right)}_{\text{Central Moment of order } n} =$$

گشتاور مرکزی مرتبه n

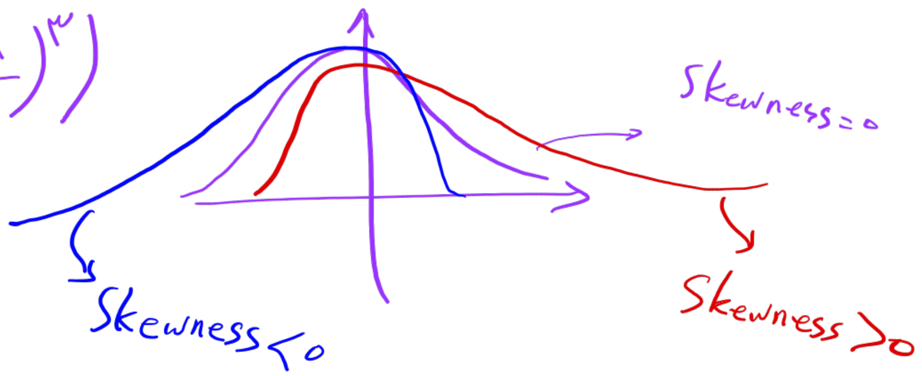
گشتاور مرکزی مرتبه ۲ : $E\left(\left(X-E(X)\right)^2\right) = Var(X)$

$$E\left(\left(\frac{X-E(X)}{\sqrt{Var(X)}}\right)^n\right) = E\left[\left(\frac{X-\mu}{\sigma}\right)^n\right] \rightarrow \text{گشتاور استاندارد مرتبه } n$$

$$\mu = E(X) \qquad \sigma = \sqrt{Var(X)}$$

گشتاور استاندارد مرتبه ۳ = Skewness چولگی

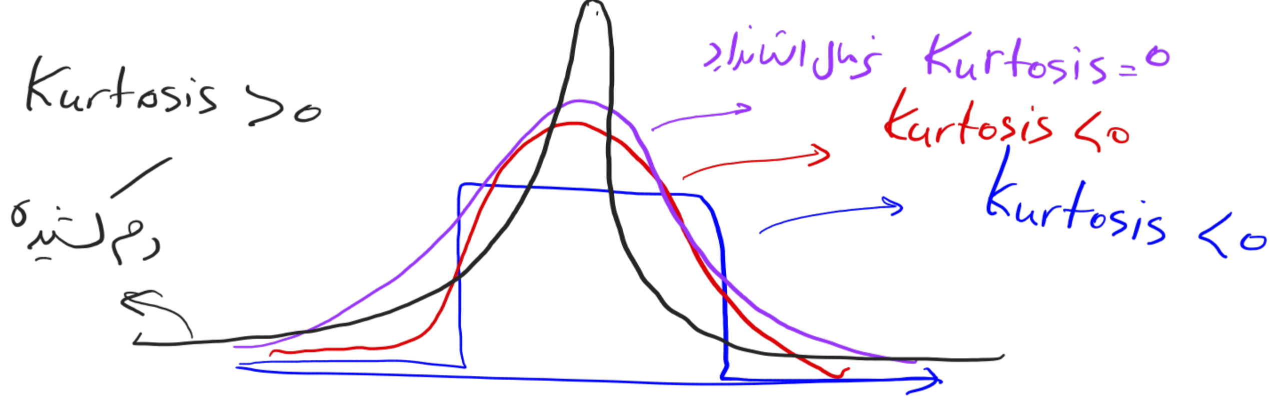
$$E\left(\left(\frac{X-\mu}{\sigma}\right)^3\right)$$



گشتاور استاندارد مرتبه ۴

Kurtosis (کشیدگی) $= E\left(\left(\frac{X-\mu}{\sigma}\right)^4\right) - 3$

توزیع نرمال استاندارد



سوال: چگونه Momentهای مرتبه بالاتر یک توزیع دلخواه را حساب کنیم؟

راه سخت: $E(X^3)=\int_{-\infty}^{+\infty} x^3 f_X(x)\,dx$

پیوسته

$$E(X^3)=\sum_{-\infty}^{+\infty} k^3 P(X=k)=\sum_{-\infty}^{+\infty} k^3 f_X(k)$$

راه زیبا: استفاده از تابع مولد گشتاور Moment Generating Function

MGF:

$$M(t)=E(e^{tX})$$

مثال: تابع مولد گشتاور توزیع برنولی:

$$M(t)=E(e^{tX})=\sum_{k=0}^{1} e^{tk}P(X=k)=pe^t+(1-p)e^0=pe^t+q$$

سوال: $M(0)=?$

$$M(0)=E(e^{0X})=1$$

همیشه بعد محاسبه MGF می‌توانیم برای $t=0$ آن را چک کنیم که مقدارش ۱ باشد.

$t=0 \to pe^{0}+q = p+q = 1$ ✓

سوال: MGF برای توزیع Binomial:

$X_i \sim Bern(p)$

$Y = X_1+X_2+\dots+X_n$

MGF $Y=X_1+X_2$?

MGF

$M_Y(t) = E(e^{tY}) = E(e^{t(X_1+X_2)})$

$= E(e^{tX_1} \times e^{tX_2}) = E(e^{tX_1})E(e^{tX_2})$

اگر X_1 و X_2 مستقل باشند

$= M_{X_1}(t)\,M_{X_2}(t)$

اگر X_1 و X_2 مستقل باشند $E(X_1X_2) = E(X_1)E(X_2)$

$M_Y(t) = M_{X_1}(t)\,M_{X_2}(t)\dots M_{X_n}(t)$

$Y = X_1+X_2+\dots+X_n$

X_i ها مستقل باشند

برنولی: $M_{X_i}(t) = pe^{t}+q$

دوجمله‌ای: $\underbrace{(pe^{t}+q)(pe^{t}+q)\dots(pe^{t}+q)}_{n\ \text{تا}} = (pe^{t}+q)^{n}$

توزیع نمایی:

$X \sim \text{Exponential}(\lambda) \qquad f_X(x) = \lambda e^{-\lambda x}$

$$M_X(t) = E(e^{tX}) = \int_{-\infty}^{+\infty} e^{tx} f_X(x)\,dx = \int_0^{+\infty} e^{tx} \lambda e^{-\lambda x}\,dx = \lambda \int_0^{+\infty} e^{x(t-\lambda)}\,dx$$

$$\frac{\lambda}{t-\lambda} e^{x(t-\lambda)} \Big|_0^{+\infty} = \frac{\lambda}{\lambda - t}$$

$t < \lambda$ فرض کنیم

سوال: از MGF چه استفاده‌ای می‌توانیم بکنیم؟

$$\underbrace{E(X^n)}_{\text{گشتاور خام مرتبه } n \text{ ام } X} = M^{(n)}(0) \quad ?$$

مثال: برنولی

$$M(t) = pe^t + q$$

$M'(t) = pe^t \qquad M'(0) = p = E(X) \qquad X \sim \text{Bern}(p)$

$M''(t) = pe^t \qquad M''(0) = p = E(X^2)$

واریانس برنولی

$$\text{Var}(X) = \underset{E(X^2) - E(X)^2}{M''(0) - M'(0)^2} = p^2 - p = p(1-p) = pq$$

دوجمله‌ای

$$M(t) = (pe^t+q)^n$$

$$M'(t) = \underbrace{npe^t}\underbrace{(pe^t+q)}^{n-1} \qquad M'(0) = np = E(X)$$

$X \sim Binomial(p, n)$

$$M''(t) = np\left[\underbrace{\underbrace{e^t}_{1}\underbrace{(pe^t+q)}_{1}^{n-1}} + \underbrace{e^t}_{1}\underbrace{pe^t}_{1}(n-1)\underbrace{(pe^t+q)}_{1}^{n-2}\right]$$

$$M''(0) = np(1+p(n-1)) = E(X^2)$$

$$Var(X) = E(X^2) - E(X)^2 = M''(0) - M'(0)^2 = np(1+p) - (np)^2 =$$

$$np(1+p(n-1) - np) = \underbrace{npq}$$

$$\underbrace{1+\cancel{np}-p}_{1-p=q}-\cancel{np}$$

اثبات

$$\underbrace{E(X^n)}_{\text{گشتاور خام مرتبه } n \text{ ام}} = M^{(n)}(0)$$

$$M(t) = M(0) + M'(0)(t) + \frac{M''(0)t^2}{2} + \cdots = \sum_{k=0}^{\infty} \frac{M^{(k)}(0)\, t^k}{k!}$$

$$M(t) = E(e^{tX}) = E\left(\sum_{k=0}^{\infty} \frac{(tx)^k}{k!}\right) = E\left(\sum_{k=0}^{\infty} x^k \frac{t^k}{k!}\right) = \sum_{k=0}^{\infty} E(X^k)\frac{t^k}{k!}$$

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

$$\sum_{k=0}^{\infty} M^{(k)}(0)\frac{t^k}{k!} = \sum_{k=0}^{\infty} E(X^k)\frac{t^k}{k!} \rightarrow M^{(k)}(0) = E(X^k) \quad \forall_{0\leq k}$$

Joint Distribution توزیع توام

در حالت گسسته: X و Y را احتمال $P(X=x, Y=y)$ را به ازای مقادیر مختلف

همه در یک جدول نمایش داده‌ایم:

X \ Y	-1	0	1
1	0.2	0.2	0.1
2	0.1	0.3	0.2
3	0	0	0

$\rightarrow$ Probability Mass Function

PMF

توزیع توام

مجموع مقادیر کل جدول باید ۱ باشد.

$$P(X=2, Y=0) = P(X=2 \cap Y=0) = P(X=2 \cap Y=0) = 0.3$$

سکه اول \	رو	پشت	سکه دوم
رو	$\frac{2}{100}$	$\frac{18}{100}$	$\frac{20}{100} = \frac{1}{5}$ رو
پشت	$\frac{25}{100}$	$\frac{55}{100}$	$\frac{80}{100} = \frac{4}{5}$ پشت
سکه اول احتمال	$\frac{27}{100}$ رو	$\frac{73}{100}$ پشت	

$\rightarrow$ احتمال حاشیه‌ای

Marginal

مثال:

این دو سکه از هم مستقل نیستند!!

در حالت مستقل با ضرب توزیع‌های حاشیه‌ای PMF این است

$\frac{27}{500}$	$\frac{73}{500}$
$\frac{108}{500}$	$\frac{292}{500}$

اگر PMF دو متغیر تصادفی با توزیع توام Joint Dist. را داشته باشیم
می توانیم براساس آن احتمال جداگانه هر یک از دو متغیر را محاسبه کنیم:

$$P(X=x) = \sum_{y} P(X=x, Y=y)$$

توزیع حاشیه‌ای
Marginal Distribution

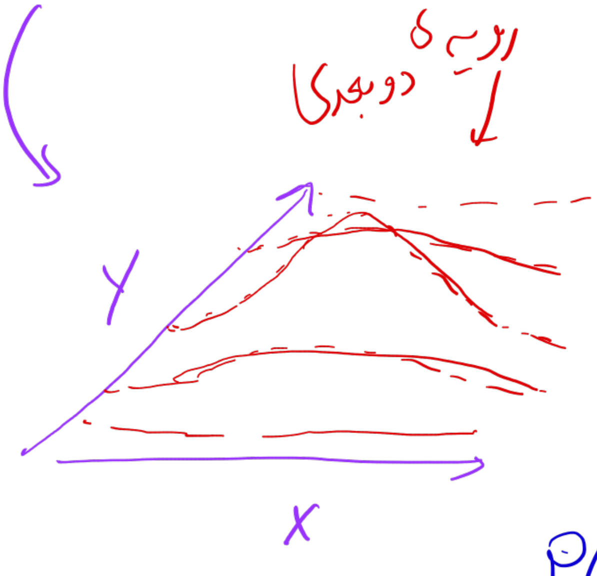
در چه حالت دو متغیر تصادفی با توزیع توام داده شده مستقل از یکدیگر هستند؟

$$\underbrace{P(X=x, Y=y)}_{\text{توزیع توام}} = \underbrace{P(X=x)}_{\text{توزیع حاشیه‌ای } X} \underbrace{P(Y=y)}_{\text{توزیع حاشیه‌ای } Y} \quad \forall x, y$$

در حالت پیوسته: PMF نداریم: به جای آن PDF توام داریم:
و CDF

$$\underbrace{f_{X,Y}(x,y)}_{\text{PDF}}$$

$$\underbrace{F_{X,Y}(x,y)}_{\text{CDF}} = P(X \leq x, Y \leq y)$$

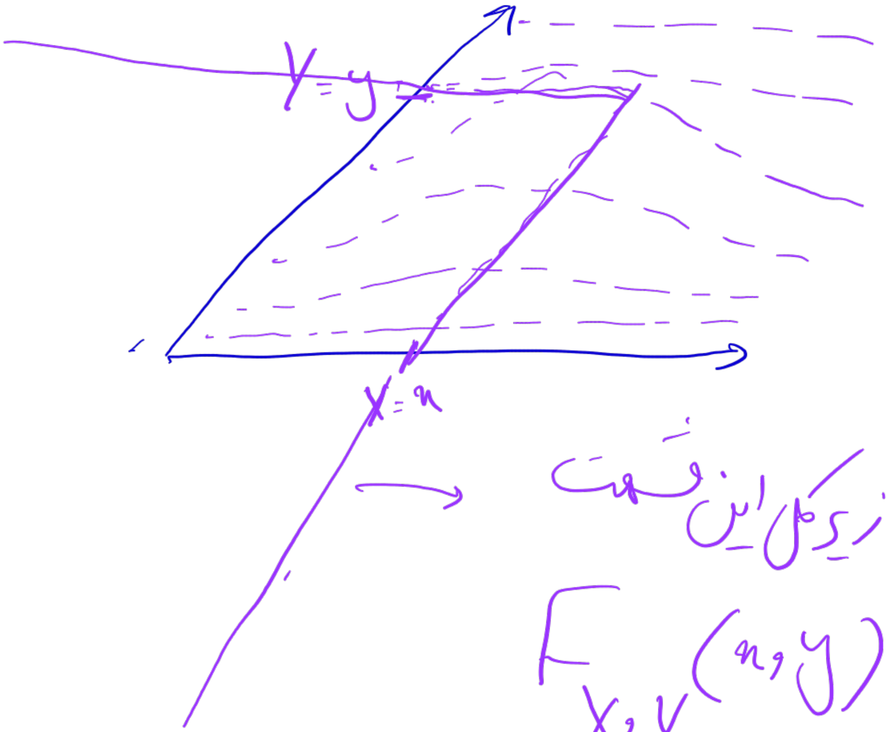


$$\int_{x=-\infty}^{+\infty} \int_{y=-\infty}^{+\infty} f_{X,Y}(x,y)\, dy\, dx = 1$$

$$P(a \leq X \leq b, c \leq Y \leq d) = \int_a^b \int_c^d f_{X,Y}(x,y)\, dy\, dx$$

$$F_{X,Y}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(u,v)\,dv\,dy$$

f مشتق F است:



حجم زیر کل این قسمت

$F_{X,Y}(x,y)$

در حالت پیوسته دو متغیر تصادفی با توزیع توام مستقل هستند اگر و تنها اگر:
به ازای هر x و y:

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

$$F_{X,Y}(x,y) = F_X(x)\,F_Y(y)$$

$$f_{X,Y}(x,y) = f_X(x)\,f_Y(y)$$

از $f_{X,Y}(x,y)$ چطور مقدار $f_X(x)$ و $f_Y(y)$ را بدست آوریم؟

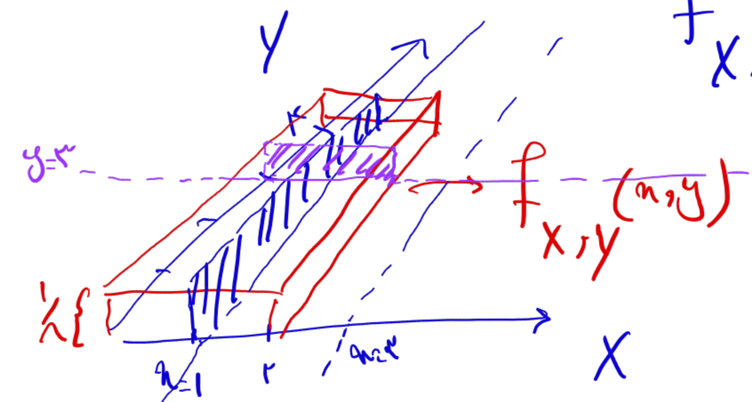
راه حل: توزیع حاشیه ای Marginal

$$f_X(x) = \int_{y=-\infty}^{+\infty} f_{X,Y}(x,y)\,dy$$

$$f_Y(y) = \int_{x=-\infty}^{+\infty} f_{X,Y}(x,y)\,dx$$

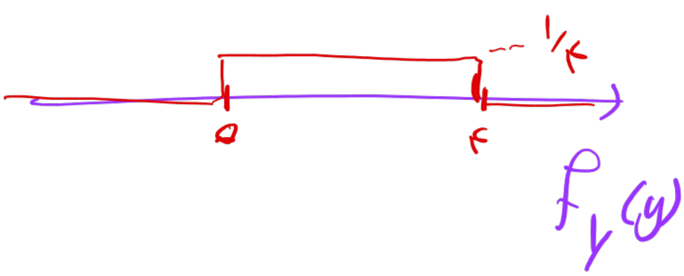
مثال: فرض کنید

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{8} & 0 \le x \le 2,\ 0 \le y \le 4 \\ 0 & \text{در غیر این صورت} \end{cases}$$



$$f_X(x) = \int_{y=-\infty}^{+\infty} f_{X,Y}(x,y)\,dy = \begin{cases} 1/2 & 0 \le x \le 2 \\ 0 & \text{در غیر این صورت} \end{cases}$$

$$f_Y(y) = \begin{cases} 1/4 & 0 \le y \le 4 \\ 0 & \text{در غیر این صورت} \end{cases}$$



مثال: یک مرغ N تخم می‌کند $N \sim \text{Poisson}(\lambda)$

و روی آن‌ها می‌نشیند. با احتمال p هر تخم جوجه می‌شود و با احتمال $q = 1 - p$ جوجه نمی‌شود.

اگر X متغیر تصادفی تعداد تخم مرغ‌هایی باشد که جوجه می‌شوند و Y تعداد آن‌هایی که جوجه نمی‌شوند

توزیع توأم PMF $(X+Y=N)$ X, Y را بدست آورید.

$$P(X=x, Y=y) = P(X=x, N=x+y) = \underbrace{P(X=x \mid N=x+y)}_{1} \underbrace{P(N=x+y)}_{2}$$

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

۱ $$P(X=x \mid N=x+y) = \binom{x+y}{x} p^x q^y$$

۲ $$P(N=x+y) = \frac{e^{-\lambda} \lambda^{x+y}}{(x+y)!}$$

$$P(X=x, Y=y) = \binom{x+y}{x} p^x q^y \frac{e^{-\lambda} \lambda^{x+y}}{(x+y)!}$$

$$= \frac{\cancel{(x+y)!}\, p^x q^y e^{-\lambda} \lambda^x \lambda^y}{x!\, y!\, \cancel{(x+y)!}} = \underbrace{\frac{e^{-\lambda p} (\lambda p)^x}{x!}}_{1} \underbrace{\frac{e^{-\lambda q} (\lambda q)^y}{y!}}_{2}$$

$$p + q = 1 \rightarrow e^{-\lambda} = e^{-\lambda p} e^{-\lambda q}$$

① → متغیر پواسون $A \sim Poisson(\lambda p)$

② $B \sim Poisson(\lambda q)$

A و B مستقل از یکدیگر هستند.

$$Var(X) = E(X^2) - E(X)^2$$

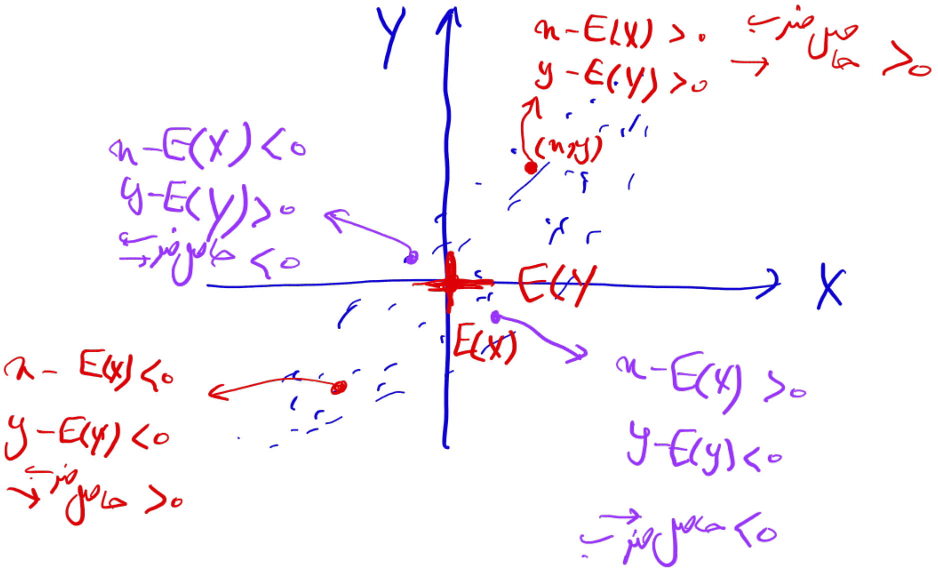
$$Var(aX+b) = a^2 Var(X)$$

$a, b \in \mathbb{R}$

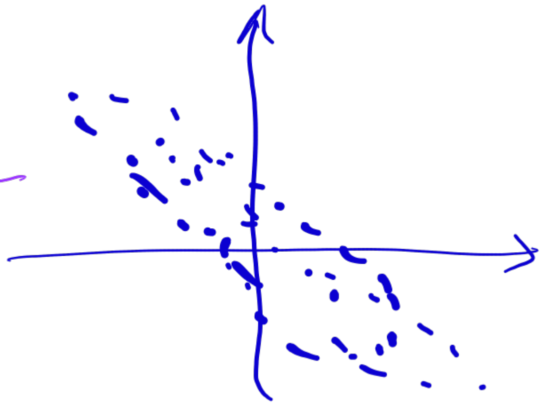
X, Y $$Cov(X,Y) = E\big((X-E(X))(Y-E(Y))\big)$$

تعداد نقاط با حاصل ضرب مثبت > نقاط با حاصل ضرب منفی

$Cov(X,Y) > 0$

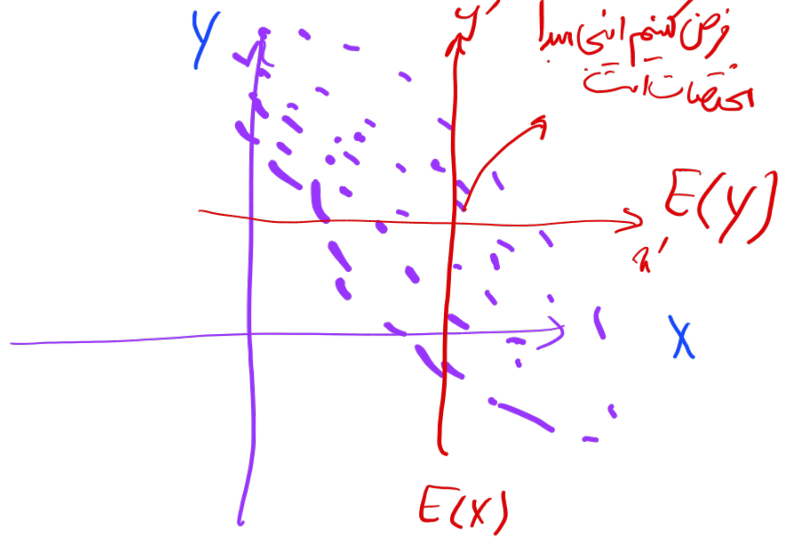


$Cov(X,Y) < 0$ ←



$x' = x - E(X)$

$y' = y - E(Y)$



چون تعداد نقاط با $x'y'<0$ بیشتر از تعداد نقاط با $x'y'>0$ است پس

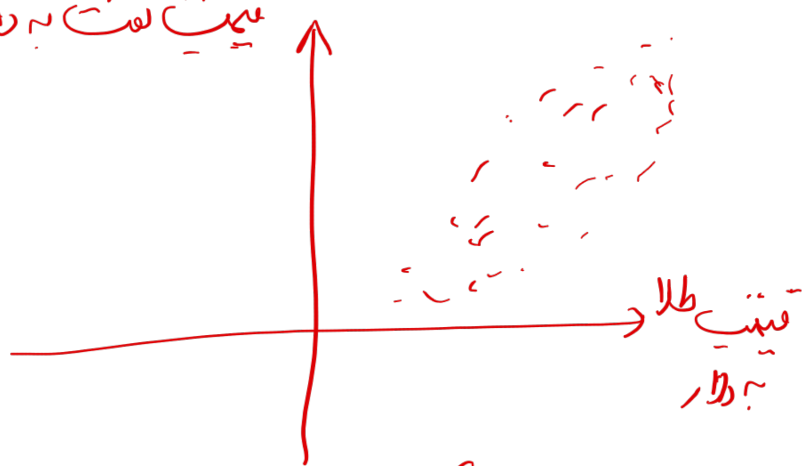
$Cov(X,Y)<0$

بیان ساده $Cov(X,Y)>0$ اگر افزایش X معمولا همراه با افزایش Y باشد.

" " کاهش X " " " کاهش Y " " .

مثلاً قد و وزن Cov مثبت دارند.

یا مثلاً وزن یک خودرو و میزان شتاب (صفر تا صد) Cov منفی دارند.



اگر دو متغیر تصادفی X,Y مستقل باشند: $Cov(X,Y)=0$

$$Cov(X,Y) = E\big[(X-EX)(Y-EY)\big]$$

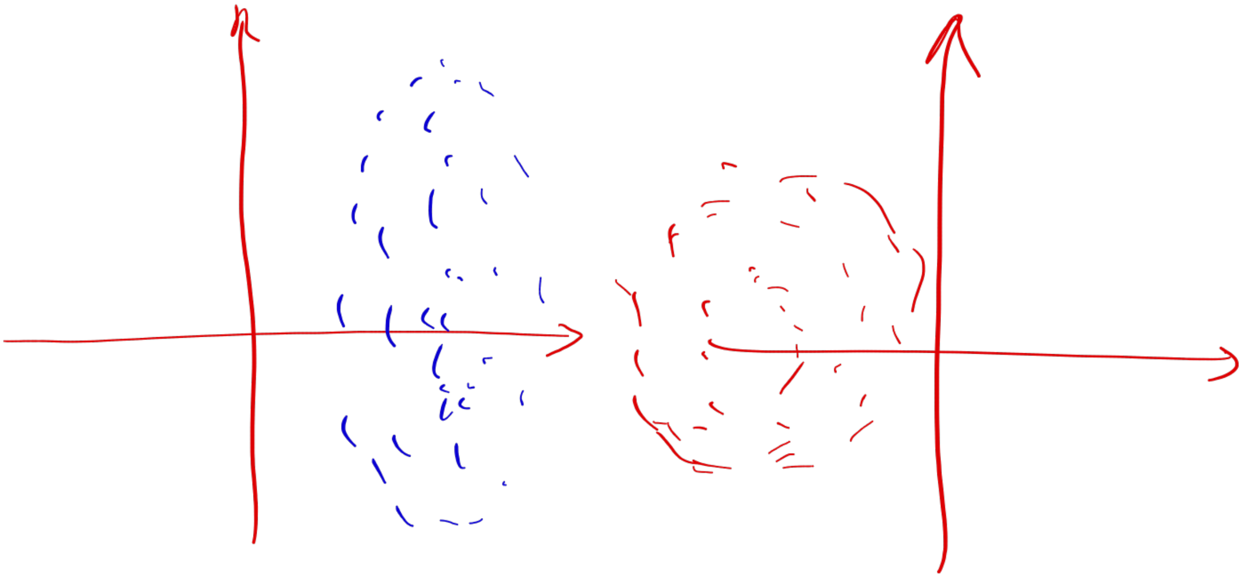
$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (x-EX)(y-EY)\, \underbrace{f_{XY}(x,y)}_{= f_X(x) f_Y(y)}\, dx\,dy = \underbrace{\int \overbrace{(x-EX)}^{\text{عدد ثابت}} f_X(x)\,dx}_{} \; \underbrace{\int \overbrace{(y-EY)}^{\text{عدد ثابت}} f_Y(y)\,dy}_{0}$$

2D LOTUS

چون X,Y مستقل هستند $= f_X(x) f_Y(y)$

$$\underbrace{\int x f_X(x)\,dx}_{E(X)} - EX \underbrace{\int f_X(x)\,dx}_{1} = 0$$

2D LOTUS : $E(g(X,Y)) = \iint g(x,y)\, f_{X,Y}(x,y)\, dx\, dy$



$Cov(X,Y) = 0$

در حالت کلی (هم X و Y مستقل باشند یا نباشند):

$$Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$$

$$Var(X-Y) = Var(X+Z) = Var(X) + \underline{Var(Z=-Y)} + 2Cov(X,-Y)$$

$Z = -Y$

$(-1)^2 = 1 \rightarrow Var(-Y) = Var(Y)$

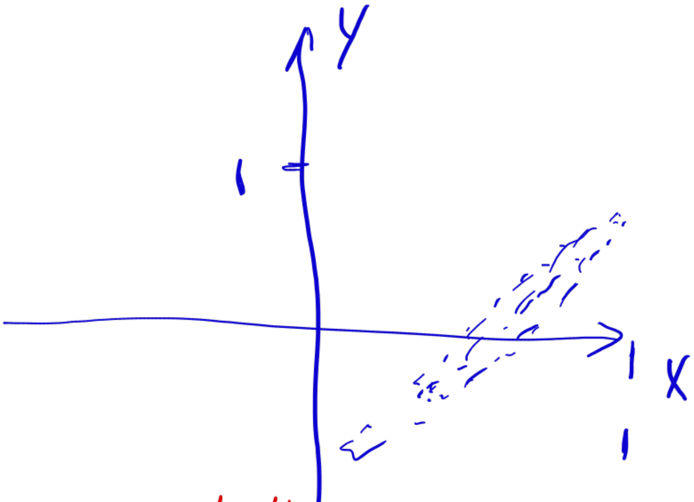
$$Cov(X,-Y) = E\Big[(X-E(X))\underbrace{(-Y-\underbrace{E(-Y)}_{-E(Y)})}_{-(Y-E(Y))}\Big] = -E((X-EX)(Y-EY)) = -Cov(X,Y)$$

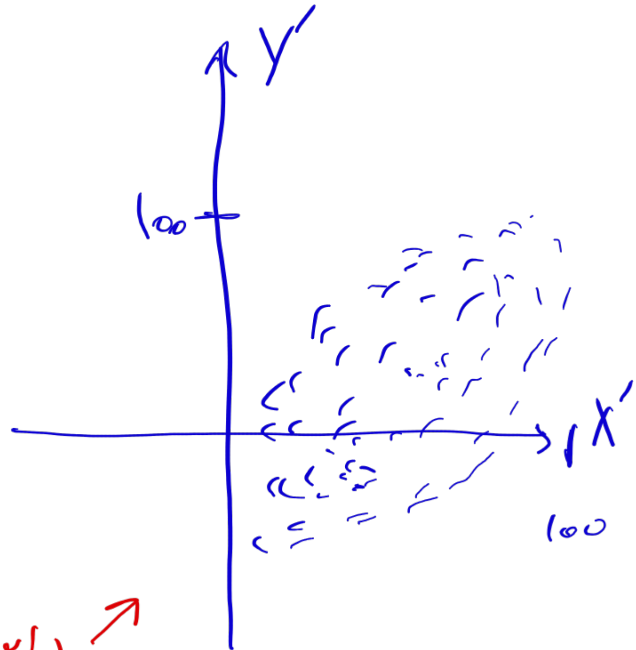
$$\Rightarrow Var(X-Y) = Var(X) + Var(Y) - 2Cov(X,Y)$$

$$Cov(aX,Y) = E\big((aX-\underbrace{E(aX)}_{aE(X)})(Y-EY)\big) = aE[(X-EX)(Y-EY)]$$

$a \in \mathbb{R}$

$$\underbrace{\qquad}_{a(X-EX)} \qquad = a\,Cov(X,Y)$$





$Cov(X',Y') \gg Cov(X,Y)$ در این شکل

$Corr(X',Y') < Corr(X,Y)$

$Cov(X,Y)$ $\qquad$ $Cov(X',Y')$

$$Corr(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}} = \frac{Cov(X,Y)}{SD(X)SD(Y)}$$

SD: Standard Deviation

قضیه: $\quad -1 \leq Corr(X,Y) \leq 1$

لم ۱: $\quad Corr(aX,Y) = Corr(X,Y)$:

$$Corr(aX,Y) = \frac{Cov(aX,Y)}{\sqrt{\underbrace{Var(aX)}_{a^2Var(X)}Var(Y)}} = \frac{aCov(X,Y)}{a\sqrt{Var(X)Var(Y)}} = Corr(X,Y)$$

اثبات قضیه

$Var(Y)=1 \quad , Var(X)=1$ بدون اینکه از کلیت قضیه کم شود فرض می‌کنم

$$Corr(X,Y) = \frac{Cov(X,Y)}{\sqrt{1 \times 1}} = Cov(X,Y) = \rho$$

$$Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y) = 1+1+2\rho = 2(1+\rho)$$

$$Var(X-Y) = Var(X) + Var(Y) - 2Cov(X,Y) = 1+1-2\rho = 2(1-\rho)$$

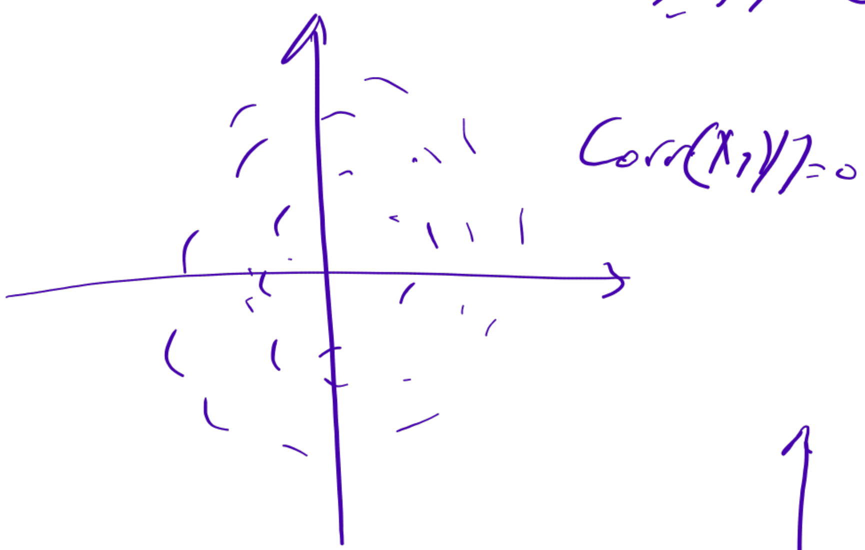
$$Var(X+Y) \geq 0 \rightarrow 1+\rho \geq 0$$

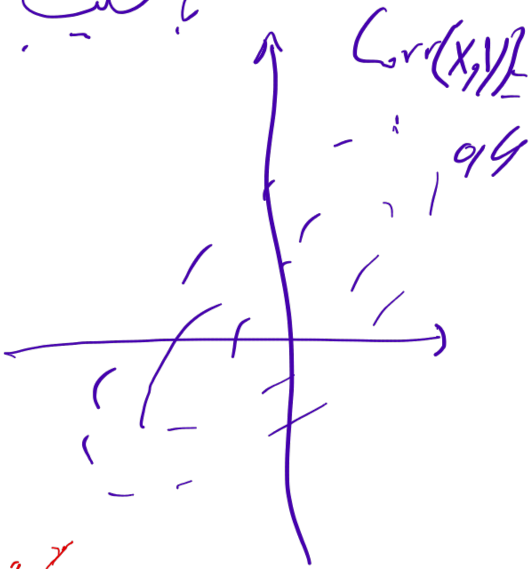
$$Var(X-Y) \geq 0 \rightarrow 1-\rho \geq 0$$

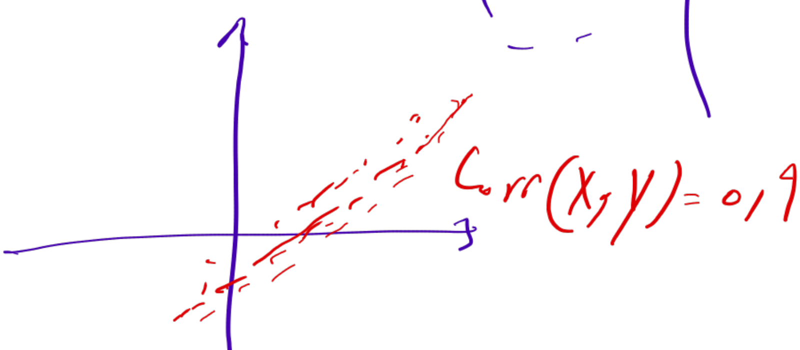
$$\rightarrow -1 \leq \rho \leq 1$$

می‌توان اثبات کرد که تنها در صورتی $Corr(X,Y)=1$ است که تمام نقاط X و Y روی یک خط با شیب مثبت قرار گیرند.

تنها در صورتی $Corr(X,Y)=-1$ است که روی خطی با شیب منفی قرار گیرند.







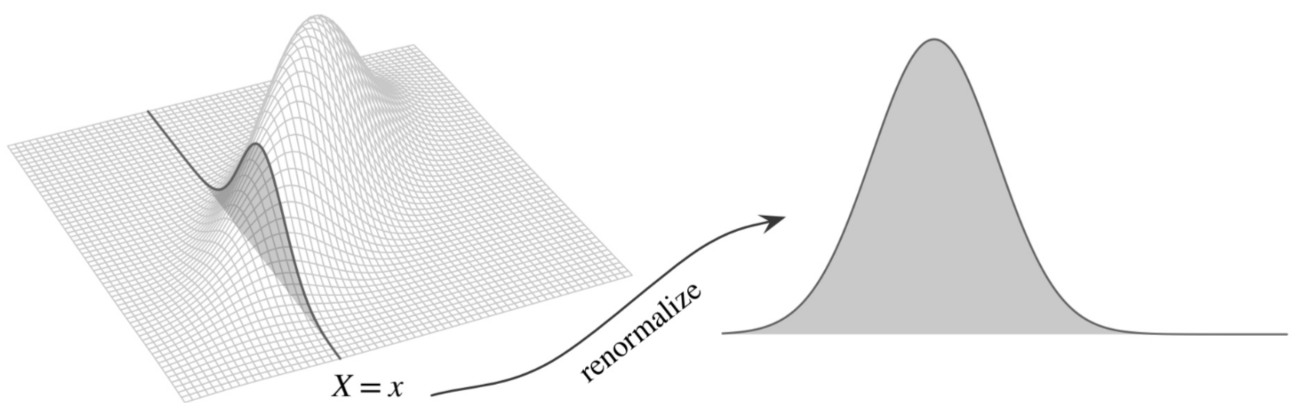
تعریف: فرض کنید چگالی توأم دو متغیر X و Y را داریم: $f_{X,Y}(x,y)$

آنگاه چگالی شرطی $X|Y$ به این صورت تعریف می‌شود:

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{f_{X,Y}(x,y)}{\int_{x=-\infty}^{+\infty} f_{X,Y}(x,y)\,dx}$$

مارجینال روی y

$$P(X|Y) = \frac{P(X,Y)}{P(Y)}$$



نابرابری های مهم احتمالاتی (فصل ۱۰ کتاب)

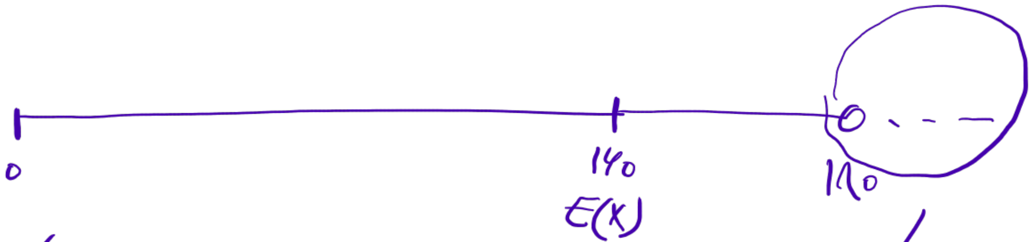
نابرابری مارکوف Markov Inequality:

به ازای متغیر تصادفی X و $\forall a>0$

$$P(|X|> a) \leq \frac{E(|X|)}{a}$$

مثال: X قد افراد است. $E(X)=160\ cm$.

$$P(X\geq 180)\leq \frac{160}{180}$$



اگر جرم بزرگی اینجا باشد $E(X)$ را به زور جابجا می کند.

چون $X\geq 0$ ، اگر مثلاً ۹۰٪ جمعیت قد ۱۸۰ به بالا داشته باشند، حتی اگر قد بقیه جمعیت صفر باشد هم:

$$E(X)\geq 0.9\times 180=162>160$$

پس اگر متوسط قد جمعیت ۱۶۰ باشد، امکان ندارد ۹۰٪ جمعیت قد ۱۸۰ به بالا داشته باشند.

اثبات Markov: ابتدا متغیر تصادفی Y را برابر $\underbrace{\frac{|X|}{a}}_{\text{عددی بزرگتر یا مساوی صفر}}$ در نظر بگیرید.

یک Indicator Variable در نظر می گیریم:

$$I(Y\geq 1)=\begin{cases}0 & Y<1 \to \frac{|X|}{a}<1\\ 1 & Y\geq 1 \to \frac{|X|}{a}\geq 1\end{cases}$$

ادعا: $I(Y\geq 1)\leq Y$

$$I(Y\geq 1)=\begin{cases}0 \rightarrow & Y\geq 0 \rightarrow \underbrace{I(Y\geq 1)}_{\text{صفر}}\leq \underbrace{Y}_{0\leq Y} \\ 1 \rightarrow & \text{این اتفاق تنها زمانی می‌افتد که } Y\geq 1 \rightarrow \underbrace{I(Y\geq 1)}_{1}\leq \underbrace{Y}_{\geq 1}\end{cases}$$

اثبات:

فقط دو مقدار دارد

$\hookrightarrow I(Y\geq 1)\leq Y$

حال از طرفین Expected می‌گیریم:

$$\underbrace{E(I(Y\geq 1))}_{\text{Indicator Variable} \text{ Expected برابر } P(I=1) \text{ است.}} = P(I(Y\geq 1)=1)=\underbrace{P(Y\geq 1)\leq E(Y)}$$

$$P\left(\frac{|X|}{a}\geq 1\right)\leq E\left(\frac{|X|}{a}\right)=\frac{E|X|}{a}$$

$$\hookrightarrow P(|X|\geq a)\leq \frac{E|X|}{a}$$

نابرابری Markov خیلی tight نیست چون فقط از $E(|X|)$ استفاده می‌کند. آیا می‌توانیم با استفاده از $Var(X)$ نابرابری محکم‌تری پیدا کنیم؟

نابرابری Chebyshev چبیشف

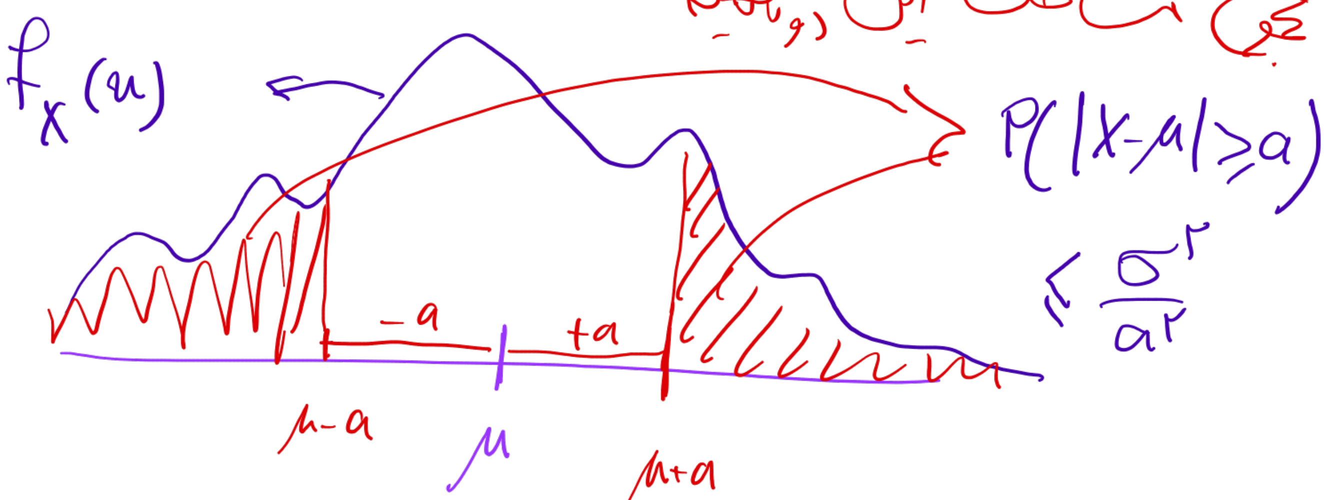
$$P(|X-\mu| \geq a) \leq \frac{\sigma^2}{a^2}$$

به ازای متغیر تصادفی X، $E(X)=\mu$ و $Var(X)=\sigma^2$ داریم:

مثال: اگر بدانیم $E(X)=140$ cm و $Var(X)=15$ آنگاه:

$$\underline{P(|X-140| \geq 20)} \leq \frac{15}{20^2} = \frac{15}{400} = \frac{3}{80}$$

$$= P(X \geq 160 \lor X \leq 120)$$



اثبات نابرابری چبیشف: از مارکوف شروع می‌کنیم:

$$P(|X| \geq a) \leq \frac{E|X|}{a}$$

طبق مارکوف

$$Y=(X-\mu)^2 \qquad P(Y \geq a^2) = P(|X-\mu| \geq a) \leq \frac{E(Y)}{a^2} = \frac{\sigma^2}{a^2}$$

$$E(Y)=E((X-EX)^2)=Var(X)=\sigma^2$$

Law of Large Numbers قانون اعداد بزرگ

بیان خیلی ساده: اگر از یک متغیر تصادفی X تعداد زیادی نمونه (مستقل) بگیریم، متوسط نمونه به سمت $E(X)$ میل می‌کند.

مثلاً اگر یک تاس را بارها (منصفانه) بیندازیم متوسط اعداد خوانده شده به ۳٫۵ میل می‌کند.

به عبارت دقیق‌تر: اگر $X_1, \ldots, X_n$ از یک توزیع یکسان و مستقل باشند (i.i.d)

آنگاه

$$\lim_{n\to\infty} \bar{X} = \lim_{n\to\infty} \frac{X_1+X_2+\cdots+X_n}{n} = \underbrace{E(X)}_{\text{امید ریاضی توزیع یا همان } E(X_i)}$$

اثبات:

$$E(\bar{X}) = E\left(\frac{X_1+\cdots+X_n}{n}\right) = \frac{nE(X_i)}{n} = E(X_i) = E(X)$$

$$Var(\bar{X}) = Var\left(\frac{X_1+\cdots+X_n}{n}\right) = Var\left(\frac{X_1}{n}\right) + Var\left(\frac{X_2}{n}\right) + \cdots + Var\left(\frac{X_n}{n}\right)$$

چون X_i ها مستقل هستند

$$= \frac{Var(X_1)}{n^2} + \frac{Var(X_2)}{n^2} + \cdots + \frac{Var(X_n)}{n^2} = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

طبق نابرابری چبیشف:

$$P(|\bar{X}-\mu| \geq a) \leq \frac{Var(\bar{X})}{a^2}$$

به ازای $\epsilon > 0$:

$$P(|\bar{X}-\mu| \geq \epsilon) = \frac{\sigma^2}{n\epsilon^2}$$

$$\lim_{n\to\infty} \frac{\sigma^2}{n\epsilon^2} = 0 \xrightarrow{\text{یعنی}}$$ از جهت احتمال اینکه $|\bar{X}-\mu|$ بیشتر از ϵ باشد صفر است.

به عبارت دیگر اگر تعداد نمونه خیلی زیاد باشد ($n \to \infty$) احتمال آن که $\bar{X}$ از مقدار واقعی μ بیش از ϵ خطا داشته باشد (به ازای هر $\epsilon > 0$ هر چقدر هم کوچک) به صفر میل می‌کند.

نابرابری کوشی شوارتز:

$$|E(XY)| \leq \sqrt{E(X^2)E(Y^2)}$$

اثبات:

به ازای هر مقدار t:

$$E((Y-tX)^2) \geq 0$$

اگر

$$t = \frac{E(XY)}{E(X^2)} \qquad (t \text{ عدد ثابت})$$

$$E((Y-tX)^2) = E(Y^2) + t^2E(X^2) - 2tE(YX) \geq 0$$

با جایگذاری $t = \frac{E(XY)}{E(X^2)}$

$$E(Y^2) + \frac{E(XY)^2}{E(X^2)^2}E(X^2) - 2\frac{E(XY)}{E(X^2)}E(XY) =$$

$$E(Y^2) + \frac{E(XY)}{E(X^2)}\left(\overbrace{\cancel{E(XY) - 2E(XY)}}^{-E(XY)}\right) \geq 0$$

$$\to E(Y^2) - \frac{E(XY)^2}{E(X^2)} \geq 0$$

$$E(X^2)E(Y^2) \geq E(XY)^2$$

$$\Rightarrow |E(XY)| \leq \sqrt{E(X^2)E(Y^2)}$$

مثال: اگر $E(X)=0$ و $E(Y)=0$ $\Leftarrow$

$$|E(XY)| = |E[(X-EX)(Y-EY)]| = |Cov(X,Y)|$$

داریم:

$$|Cov(X,Y)| \leq \sqrt{E(X^2)E(Y^2)} = \sqrt{Var(X)Var(Y)}$$

$$Var(X) = E(X^2) - \underbrace{E(X)^2}_{0} = E(X^2)$$

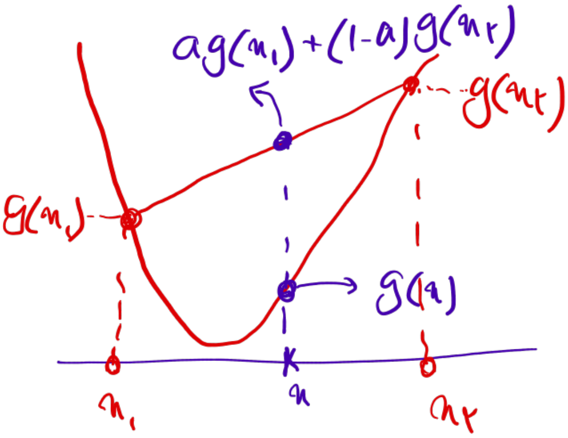
$$\frac{|Cov(X,Y)|}{\sqrt{Var(X)Var(Y)}} = |Corr(X,Y)| \leq 1$$

Jensen Inequality — نابرابری Jensen

برای هر تابع محدب $g(x)$ و هر متغیر تصادفی X داریم:

$$E(g(X)) \geq g(E(X))$$

تابع محدب : تابعی که خط واصل هر دو نقطه بالای تابع در فاصله بین آن دو نقطه قرار گیرد:



$x = a x_1 + (1-a) x_2$

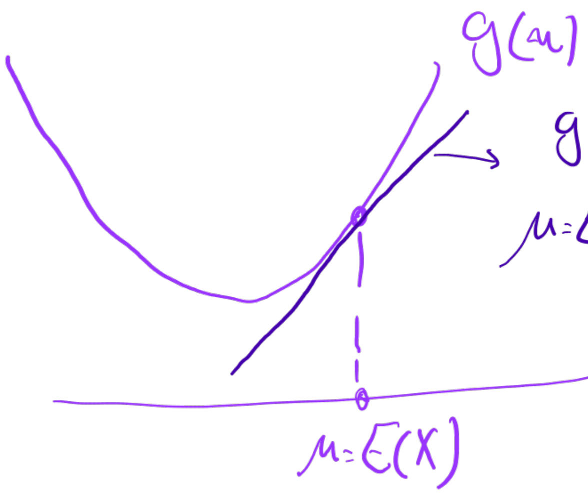
$0 \le a \le 1$

$$\begin{cases} a=1 \rightarrow x = x_1 \\ a=0 \rightarrow x = x_2 \end{cases}$$

تابع g محدب است اگر $\forall x_1, x_2$ و $0 \le a \le 1$:

$$\underbrace{a g(x_1) + (1-a) g(x_2)}_{\text{خط واصل دو نقطه } (x_1, g(x_1)) \text{ و } (x_2, g(x_2))} \ge \underbrace{g(a x_1 + (1-a) x_2)}_{\text{قسمتی از } g \text{ که بین } x_1, x_2 \text{ قرار دارد}}$$

اثبات نابرابری Jensen:



خط مماس بر g در نقطه $\mu = E(X)$ → معادله این خط را $ax+b$ در نظر بگیرید.

می‌دانیم که $g(x)$ بالای این خط $ax+b$ است

یعنی $g(x) \ge ax+b$

$$E(g(X)) \ge E(aX+b) = aE(X)+b = g(\mu) = g(EX)$$

چون خط $ax+b$ در نقطه μ به تابع g مماس است

یعنی $g(\mu) = g(EX) = a\mu + b$

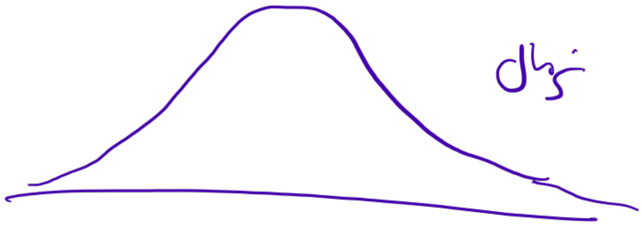
نابرابری Jensen در حالت n بعدی هم کار می‌کند.

Central Limit Theorem : قضیه حد مرکزی

اگر $X_1, X_2, \dots, X_n$ متغیرهای تصادفی مستقل از یک توزیع باشند (iid)
آنگاه مجموع یا متوسط آن‌ها در حالتی که $n \to +\infty$ برود از توزیع نرمال تبعیت می‌کند.

X

$X_1 + X_2 + \cdots + X_n \longrightarrow$



نرمال

کاربرد: ۱) فرض کنید زمان قبول سفر توسط هر راننده تاکسی اسنپی از یک توزیع نمایی با پارامتر λ تبعیت می‌کند. احتمال آنکه متوسط زمان قبول سفر توسط ۳۰ راننده از 2λ بیشتر شود چقدر است؟ با واریانس ۱۰۰ ساعت

۲) زمان عمر هر یک لامپ به طور متوسط ۱۰۰۰ ساعت است. یک کارتن لامپ شامل ۱۰۰ لامپ است. اگر متوسط زمان کارکرد لامپهای هر جعبه زیر ۹۹۷ ساعت باشد شامل گارانتی می‌شود و شرکت باید ۱۰۰۰ دلار به فروشگاهی که کارتن لامپ را خریده پرداخت کند. به طور متوسط شرکت به هر فروشگاه چقدر جریمه خواهد پرداخت؟

فرض کنید زمان کارکرد هر لامپ یک متغیر تصادفی X_i باشد.

$$X_i \sim D(\mu = 1000, \sigma^2 = 100)$$

توزیع ناشناخته

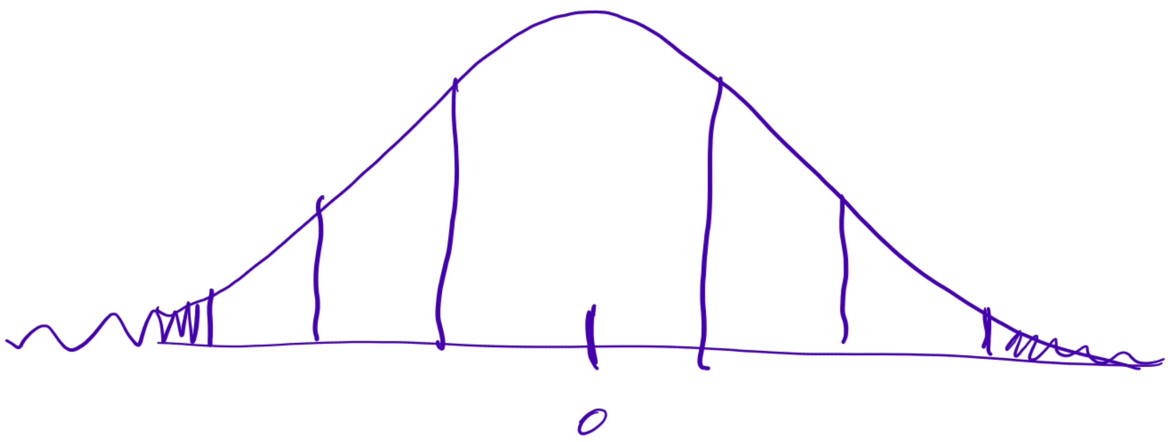
$$\bar{X} = \frac{X_1 + \cdots + X_{100}}{100} \xrightarrow[\text{قضیه حد مرکزی}]{CLT} N(\mu = 1000, \sigma^2 = 1)$$

$$\sigma^2(\bar{X}) = \underbrace{\frac{100}{10000} + \frac{100}{10000} + \cdots + \frac{100}{10000}}_{100} = 1$$

سوال: $\bar{X} \sim N(1000, 1)$

؟ $P(\bar{X} < ۹۹۷)$

پاسخ: $Y \sim N(0,1)$ ، $P(Y < -۳)$



$$P(-۳ \le Y \le ۳) = ۹۹٫۷\ \%$$

$P(Y<-۳) = ۰٫۱۵\% = ۰٫۰۰۱۵$

E(خسارت گارانتی شرکت به ازای هر وسیله) $= ۰٫۰۰۱۵ \times 1000\$ = ۱٫۵\$$

نکته عجیب: اگر مجبور به پرداخت ۱۵٪ برای متوسط زمان کارکرد زیر ۹۹۷ ساعت

که ضرر است برای متوسط زیر ۱۰۰۰ ساعت، ۱۵ درصد؟ یا Expected؟

به جای ۱٫۵$ ضرر کردن برابر ۵۰۰$ بود!

اثبات CLT :

لم ۱) اگر $X \sim N(0,1)$ آن‌گاه تابع مولد گشتاور (MGF) متغیر X برابر با $e^{t^2/2}$ است.

$$MGF(X) = M_X(t) = E(e^{tX})$$

یادآوری: $M_X^{(n)}(0)$ یا مشتق n ام M در نقطه صفر برابر گشتاور n ام X بود: $E(X^n)$

$$X \sim N(0,1) \rightarrow E(e^{tX}) \underset{\text{LOTUS}}{=} \int_{-\infty}^{+\infty} e^{tx} f_X(x)\,dx = \int_{-\infty}^{+\infty} e^{tx} \underbrace{\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}}_{f_X(x) = \varphi(x)} dx$$

$$tx - \frac{x^2}{2} = -\frac{(x-t)^2}{2} + \frac{t^2}{2} \rightarrow \int e^{tx-\frac{x^2}{2}} \frac{1}{\sqrt{2\pi}}\, dx = e^{t^2/2} \underbrace{\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-t)^2}{2}} dx}$$

اگر به جای $x-t$ ، y را بنویسیم این انتگرال برابر $\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy$ است

که برابر $\int_{-\infty}^{+\infty} f_X(x)\,dx$ برای توزیع نرمال استاندارد است ⇐ برابر ۱ است.

در واقع این انتگرال روی تابع چگالی توزیع نرمال $N(\mu=t, \sigma^2=1)$ است.

$$X \sim N(0,1) \rightarrow E(e^{tX}) = e^{t^2/2}$$

اثبات CLT:

فرض کنید $X_1 \dots X_n \sim D(\mu=0, \sigma^2=1)$ تبعیت می‌کنند که D هر توزیع دلخواهی است،

اگر میانگین و واریانس D برابر صفر و یک نبود و به جای آن برابر μ و σ^2 بود می‌توانیم

$$X_i' = \frac{X_i - \mu}{\sqrt{\sigma^2}} \rightarrow X_i' \sim D'(\mu=0, \sigma^2=1)$$

تبعیت خواهند کرد.

حالا نشان می‌دهیم اگر $X_1 \dots X_n \sim D(0,1)$ آنگاه: $Z = \frac{X_1 + \dots + X_n}{\sqrt{n}} \sim N(0,1)$

وقتی $n \to \infty$

کافی است نشان دهیم $MGF_Z(t) = e^{t^2/2}$ است.

قضیه Uniqeness: اگر دو توزیع احتمالاتی MGF برابر در یک بازه‌ی باز شامل نقطه‌ی صفر داشته باشند، همهٔ دو توزیع برابر هستند.

اثبات: فرض کنید MGF ی $D(0,1)$ وجود دارد و آن را $M(t)$ بنامیم.

$$MGF(Z)=E(e^{tZ})=E\left(e^{t\left(\frac{X_1+X_2+\cdots+X_n}{\sqrt{n}}\right)}\right)=$$

$$E\left(e^{\frac{tX_1}{\sqrt n}}\right)E\left(e^{\frac{tX_2}{\sqrt n}}\right)\cdots E\left(e^{\frac{tX_n}{\sqrt n}}\right)$$ چون X_i ها مستقل هستند.

$$=M\left(\frac{t}{\sqrt n}\right)M\left(\frac{t}{\sqrt n}\right)\cdots M\left(\frac{t}{\sqrt n}\right)=M\left(\frac{t}{\sqrt n}\right)^n$$

$$\lim_{n\to\infty} M\left(\frac{t}{\sqrt n}\right)^n=e^{\frac{t^2}{2}}$$ می‌خواهیم نشان دهیم

$\to$ از طرفین log می‌گیریم:

$$\lim_{n\to\infty} n\log M\left(\frac{t}{\sqrt n}\right)=\lim_{y\to 0}\frac{1}{y^2}\log M(ty)$$ $\to$ این کسر

اگر y را برابر $\frac{1}{\sqrt n}$ تعریف کنیم

$\frac{0}{0}$ است پس از هوپیتال استفاده می‌کنیم:

اولین مشتق MGF روی نقطه صفر برابر است با $E(X)=0$

$$\lim_{y\to 0}\frac{\log M(ty)}{y^2}=\lim_{y\to 0}\frac{t\,M'(ty)}{2y\,M(ty)}=\frac{t}{2}\lim_{y\to 0}\frac{M'(ty)}{y\,M(ty)}$$

باز هم کسر $\frac{0}{0}$ داریم دوباره هوپیتال

$$\frac{t}{2}\lim_{y\to 0}\frac{M'(ty)}{yM(ty)}=\frac{t}{2}\lim_{y\to 0}\frac{tM''(ty)}{M(ty)+\underbrace{ytM'(ty)}_{0}}=\frac{t^2}{2}\lim\frac{M''(ty)}{\underbrace{M(ty)}_{1}}$$

$$M''(0)=E(X^2)=1\quad (Var(X)=E(X^2)-\overbrace{E(X)}^{0}{}^2=1)$$

$$\rightarrow \lim_{n\to\infty} n \log M\left(\frac{t}{\sqrt{n}}\right) = \frac{t^2}{2} \rightarrow$$

$$\lim_{n\to\infty} MGF_Z = e^{\frac{t^2}{2}} \rightarrow MGF$$

توزیع نرمال استاندارد

توزیع Multinomial چندجمله‌ای

اگر یک آزمایش را n بار تکرار کنیم و در هر بار مقدار مشاهده شده یکی از مقادیر ۱ تا k باشد

$(1 \le i \le k)$ X_i را تعداد مشاهده i در n بار آزمایش در نظر بگیریم، آنگاه

$$(X_1, X_2, \dots, X_k) \sim Multinomial(n, P_1, P_2, \dots, P_k)$$

تبعیت می‌کند احتمال دیدن هر مقدار i برابر P_i باشد.
$1 \le i \le k$

$P_1 + P_2 + \dots + P_k = 1$

مثال: تاس غیر منصفانه:

فرض کنید

$P_1 = P_2 = P_3 = P_4 = P_5 = 0/1$

$P_6 = 0/5$

و X_i برابر تعداد دفعات آمدن i در n بار پرتاب تاس باشد.
$1 \le i \le 6$

$$(X_1 \dots X_6) \sim Multi(n, P_1 \dots P_6)$$

$n_1 + \dots + n_k = n$

$$PMF: \quad P(\underbrace{X_1 = n_1}_{n_1 \text{ بار ۱ بیاید}}, X_2 = n_2, \dots, X_k = n_k)$$

$\underline{2}\ \underline{3}\ \underline{1}\ \underline{2}$ ... $\underline{\ \ }\ \underline{\ \ }\ \underline{\ \ }$

n بار

$$P(X_1=n_1, X_2=n_2 \cdots X_k=n_k) = \binom{n_1+\cdots+n_k}{n_1}\binom{n_2+\cdots+n_k}{n_2}\binom{n_3+\cdots+n_k}{n_3}\cdots\binom{n_k}{n_k} p_1^{n_1} p_2^{n_2}\cdots p_k^{n_k}$$

$$= \frac{(n_1+\cdots+n_k)!}{n_1!\,(n_2+\cdots+n_k)!}\cdot\frac{(n_2+\cdots+n_k)!}{n_2!\,(n_3+\cdots+n_k)!}\cdot\frac{(n_3+\cdots+n_k)!}{n_3!\,(n_4+\cdots+n_k)!}\cdots\frac{n_k!}{n_k!\,0!}\, p_1^{n_1} p_2^{n_2}\cdots p_k^{n_k}$$

$$= \frac{(n_1+n_2+\cdots+n_k)!}{n_1!\, n_2! \cdots n_k!}\, p_1^{n_1} p_2^{n_2}\cdots p_k^{n_k}$$

اگر $(X_1 \cdots X_k) \sim Multinomial(n, p_1 \cdots p_k)$

$E(X_1)?\quad E(X_2)? \quad \cdots \quad E(X_k)?$

$E(X_1)=np_1 \quad E(X_2)=np_2 \quad \cdots \quad E(X_k)=np_k$

مثال: اگر تاس بالا را ده بار بیندازیم احتمال آنکه

$n_1=1,\ n_2=2,\ n_3=1,\ n_4=2,\ n_5=1,\ n_6=3$

باشد چقدر است؟

$X_1=n_1=1$

$X_2=n_2=2$

$X_3=n_3=1$

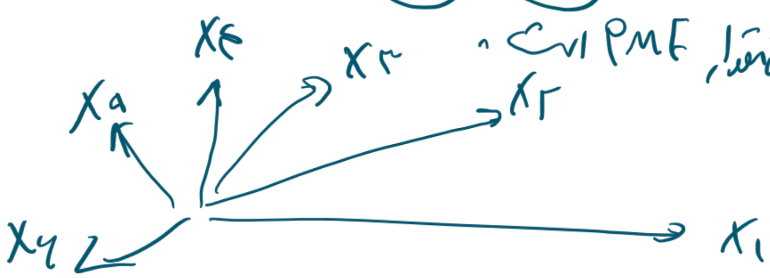
$\vdots$

$X_6=n_6=3$

$(X_1, X_2, X_3 \ldots X_6)$

فضای PMF یک فضای ۶ بعدی است

و البته یک بعد هفتم که مقدار PMF است.



احتمال $(1,2,1,2,1,3)$ چقدر است؟

$$\frac{10!}{1!\,2!\,1!\,1!\,2!\,1!\,3!}(0.1)^1(0.1)^2(0.1)^1(0.2)^2(0.1)^1(0.5)^3 =$$

اثبات آنکه $E(X_1)=np_1 \quad E(X_2)=np_2 \cdots \quad E(X_k)=np_k$

دیدن هر مقدار i را می‌توانیم به صورت یک متغیر تصادفی دو حالته در نظر بگیریم

$$\begin{array}{cccccccccc} 0 & 0 & 1 & 0 & 0 & 1 & 0 & \cdots & 1 \\ \underset{1}{1} & \underset{2}{4} & 3 & 2 & 6 & 3 & 5 & & \underset{n}{3} \end{array}$$

تاس را n بار انداختیم

دفعاتی که تاس i بیاید را ۱ و سایر دفعات را صفر در نظر بگیریم.

مثال: $i=3$ احتمال آنکه تاس ۳ بیاید چقدر است؟ p_3

$$E(X) = np_3$$

$$X \sim Binom(n, p_3)$$

در نتیجه $E(X_i)=np_i$ به ازای هر $1 \le i \le k$

اگر می‌خواستم از مقدار PMF، $E(X_i)$ را حساب کنم باید این کار را می‌کردم:

$$E(X_i)=\sum_{n_i=0}^{n} n_i \underbrace{P(X_i=n_i)}_{\text{احتمال حاشیه‌ای}} = \sum_{n_i=0}^{n} \sum_{n_1,n_2,\dots,n_{i-1},n_{i+1},n_{i+2},\dots,n_k} n_i P(X_1=n_1, X_2=n_2, \dots, X_k=n_k)$$

Marginal X_i ای

$\rightarrow n_1+n_2+\dots+n_k=n$

$\rightarrow n_1+n_2+\dots+n_{i-1}+n_{i+1}+\dots+n_k=n-n_i$

در واقع : $(EX_1, EX_2, \dots EX_k)$:

آزمایش ۱ :

$(x_1, x_2, \dots, x_k)$ n بار تاس را که وجهی با احتمال های $p_1 \dots p_k$ را بیندازید و تعداد دفعات آمدن i را بشمارید و آن را x_i بنامید.

$n = 10$

تاس   1   3   2   1   4   2   6   1   3   2

$(x_1 = 3, x_2 = 3, x_3 = 2, x_4 = 1, x_5 = 0, x_6 = 1)$

آزمایش ۲: همان کار را تکرار کنید یعنی تاس را n بار دیگر بیندازید.

آزمایش ∞: همان کار را $N \Rightarrow \infty$ بار تکرار کنید

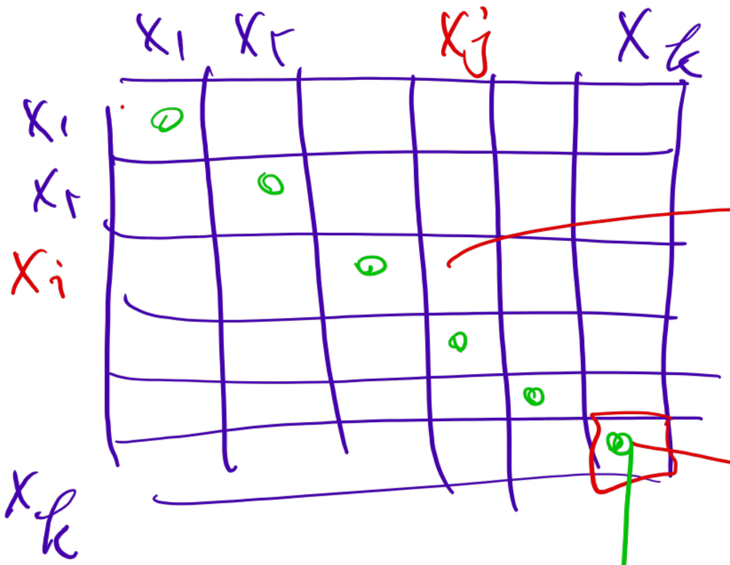
	x_1	x_2	$\dots$	x_k
آزمایش ۱				
آزمایش ۲				
$\vdots$				
آزمایش N				

$\longrightarrow$ | EX_1 | EX_2 | $\dots$ | EX_k |

روی هر ستون متوسط بگیرید

برای $(X_1 \dots X_k) \sim Multi(n, p_1 \dots p_k)$

ماتریس Variance Covariance داریم یک مقدار واریانس.



$Cov(X_i, X_j)$

$Cov(X_k, X_k) = E((X_k - EX_k)(X_k - EX_k)) = E((X_k - EX_k)^2)$

$= Var(X_k)$

پس در قطر اصلی $Var(X_i)$ ها داریم

$$Var(X_i) = E((X_i - EX_i)^2) = np_i(1-p_i)$$

اثبات: به همان استدلالی که (X_i را می‌توانیم تعداد تیک ها در یک آزمایش Binom با n بار پرتاب سکه با احتمال p_i برای آن که سکه در هر پرتاب i بیاید)

$$X \sim Binom(n, p_i) \rightarrow E(X) = np_i$$

$$Var(X) = np_i(1-p_i)$$

حالا $Cov(X_i, X_j)$ ؟

راه حل: تعریف کنیم $Y = X_i + X_j$

در واقع Y معادل یک سکه است که اگر تاس به وجهی مثل i یا j بیاید
و در غیر این صورت صفر می‌شود. سکه با احتمال $P = P_i + P_j$ یک می‌آید.

$$E(Y) = n(P_i + P_j)$$

$$Var(Y) = n(P_i+P_j)(1-P_i-P_j) \quad (1)$$

$$(2)\quad Var(Y) = Var(\underbrace{X_i + X_j}_{\text{مستقل نیستند}}) = \underbrace{Var(X_i)}_{nP_i(1-P_i)} + \underbrace{Var(X_j)}_{nP_j(1-P_j)} + 2Cov(X_i, X_j)$$

$$\Rightarrow (1),(2): \quad n(P_i+P_j)(1-P_i-P_j) = nP_i(1-P_i) + nP_j(1-P_j) + 2Cov(X_i, X_j)$$

$$\Rightarrow Cov(X_i, X_j) = \frac{n\left[P_i - P_i^2 - P_iP_j + P_j - P_iP_j - P_j^2 - P_i + P_i^2 - P_j + P_j^2\right]}{2}$$

$$= -nP_iP_j$$

قطر اصلی

ماتریس واریانس - کوواریانس

$$n\begin{array}{c} \\ \left[\begin{array}{ccc} \boxed{P_i(1-P_i)} & & \\ & \boxed{-P_iP_j} & \\ & & \end{array}\right] \end{array} \begin{array}{c} \\ i \end{array}$$

(column j)

توزیع نرمال چند متغیره Multivariate Normal

یک بردار $(X_1, X_2, \dots, X_k)$ از توزیع نرمال چند متغیره تبعیت می‌کند اگر به ازای همهٔ مقادیر $t_1, t_2, \dots, t_k \in \mathbb{R}$، مقدار $t_1X_1 + t_2X_2 + \dots + t_kX_k$ از توزیع نرمال پیروی کند.

قضیه ۱: اگر $(X_1, \dots, X_k) \sim \text{MultNormal}$ تبعیت کند هر یک از X_iها از توزیع نرمال تبعیت می‌کنند.

اثبات: $t_i = 1$ و سایر $t_j = 0$

$1 \leq j \leq k, \ j \neq i$

قضیه ۲: اگر $(X_1, \dots, X_k) \sim \text{MultNormal}$ آنگاه $(a_1X_1+b_1, a_2X_2+b_2, \dots, a_kX_k+b_k)$ از توزیع نرمال چند متغیره تبعیت می‌کند.

اثبات: $Y_1 = a_1X_1 + b_1$، $Y_2 = a_2X_2 + b_2$ و ... و $Y_k = a_kX_k + b_k$

می‌خواهیم ثابت کنیم $(Y_1, Y_2, \dots, Y_k)$ از توزیع نرمال چند متغیره تبعیت می‌کند.

به ازای هر $t_1, t_2, \dots, t_k \in \mathbb{R}$ دلخواه داریم:

$$t_1Y_1 + t_2Y_2 + \dots + t_kY_k = t_1(a_1X_1+b_1) + t_2(a_2X_2+b_2) + \dots + t_k(a_kX_k+b_k)$$

$$= \underbrace{t_1a_1}_{\text{عدد ثابت}}X_1 + \underbrace{t_2a_2}_{\text{عدد ثابت}}X_2 + \dots + \underbrace{t_ka_k}_{\text{عدد ثابت}}X_k + \underbrace{t_1b_1 + t_2b_2 + \dots + t_kb_k}_{\text{عدد ثابت}}$$

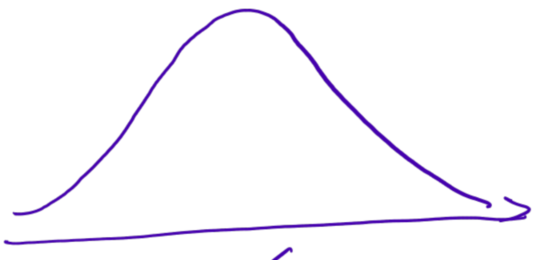
چون $(X_1, \dots, X_k) \sim \text{MultNormal}$ می‌دانیم $t_1a_1X_1 + t_2a_2X_2 + \dots + t_ka_kX_k$ باید از توزیع نرمال تبعیت کند. توزیع نرمال + عدد ثابت هم توزیع نرمال است.

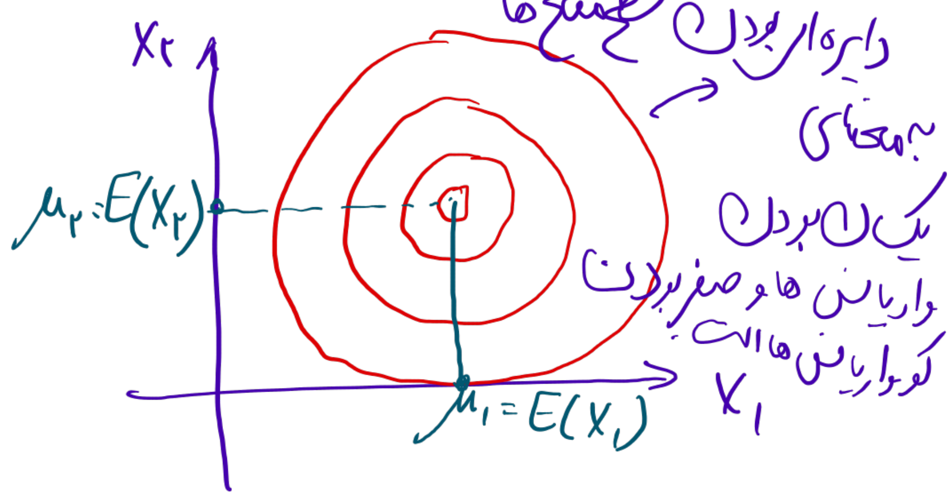
تعمیم ۳: $(X_1 \ldots X_k) \sim \text{Mult Normal}$, $(Y_1 \ldots Y_k) \sim \text{Mult Normal}$

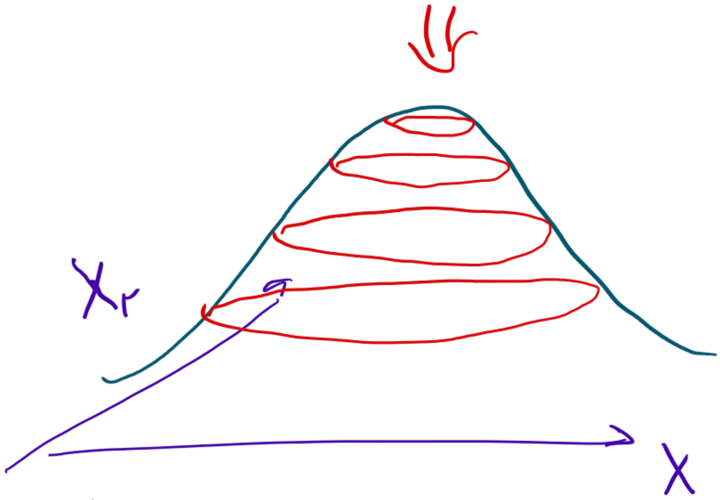
آنگاه

$$(X_1 \ldots X_k, Y_1 \ldots Y_k) \sim \text{Mult Normal}$$

شباهت به گوجه ی خودمان :D







در توزیع بالا $(X_1, X_2) \sim \text{Mult Normal}$ تبعیت می کند.

در نمودار بالا $Var(X_1) = Var(X_2)$ و $Cov(X_1, X_2) = 0$

یعنی $Corr(X_1, X_2) = 0$ یعنی X_1 و X_2 مستقل هستند.

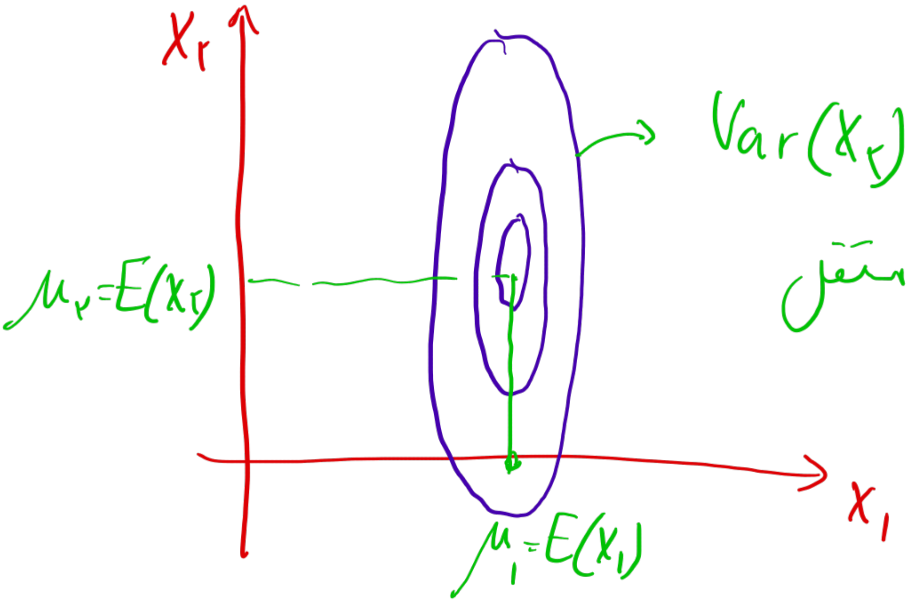
یعنی $\begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix}$ ماتریس واریانس کوواریانس X_1 و X_2 است.

یعنی: $f_{X_1,X_2}(x_1,x_2) = f_{X_1}(x_1)\, f_{X_2}(x_2) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x_1-\mu_1)^2+(x_2-\mu_2)^2}{2\sigma^2}}$

مثال ۲:

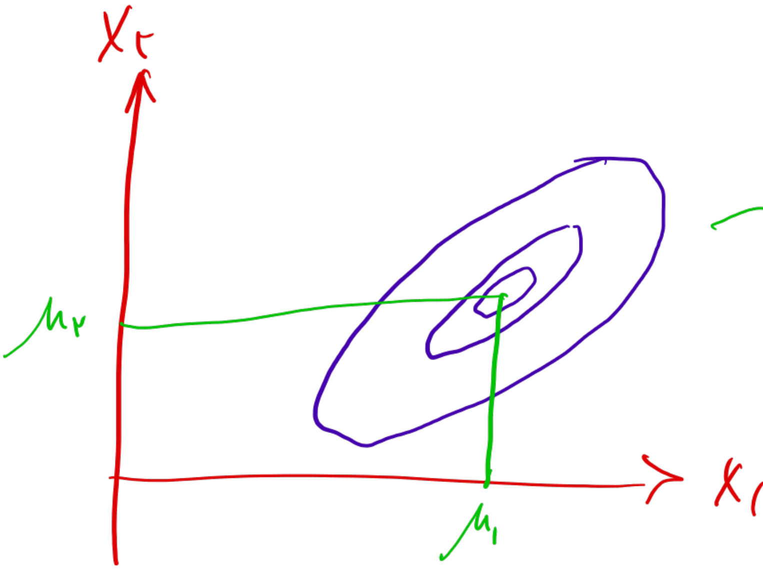
$Var(X_2) > Var(X_1)$

و X_1 و X_2 مستقل



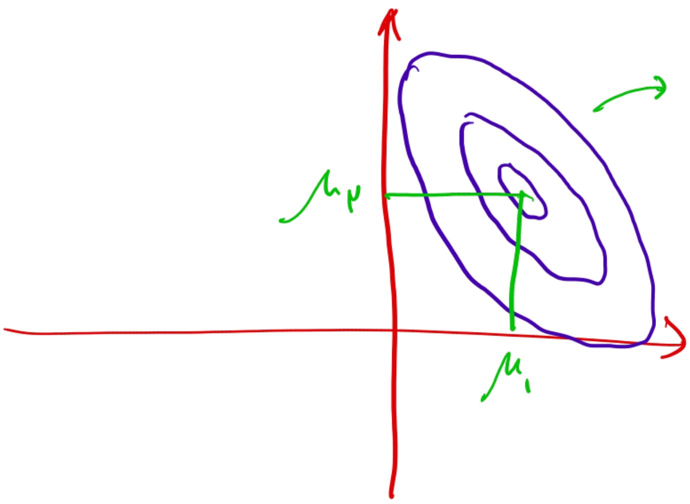
مثال ۳:

$Cov(X_1,X_2)>0$



مثال ۴:

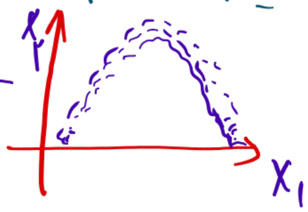
$Cov(X_1,X_2)<0$



نکته :

Correlation

• " اگر دو متغیر X_1 و X_2 کوواریانس یا همبستگی صفر داشته باشند مستقل هستند" نادرست



← مستقل نیستند ولی کوواریانس صفر دارند.

• " اگر دو متغیر X_1 و X_2 مستقل باشند کوواریانس و همبستگی صفر دارند" درست

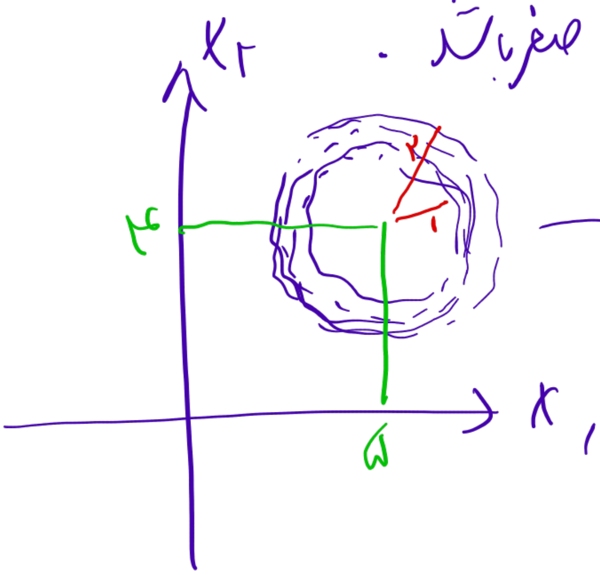
کوواریانس و همبستگی ، وابستگی خطی دو متغیر را می تواند بسنجد.

در واقع اگر دو متغیر X_1 و X_2 مستقل خطی باشند ← کوواریانس صفر دارند.

• وابسته خطی باشند ← کوواریانس غیر صفر دارند.

ولی اگر وابستگی غیرخطی داشته باشند (مثلاً $X_2 = X_1^2$)

کوواریانس آن ها ممکن است صفر باشد.



→ وابسته اند چراکه

$$1^2 \le (X_1-5)^2 + (X_2-4)^2 \le 2^2$$

ولی کوواریانس صفر دارند.

Tranfomations

فرض کنیم یک متغیر تصادفی X داریم. حال اگر $Y=g(X)$
می‌خواهیم تابع چگالی احتمال Y را به دست آوریم.

تا الان جدی‌ترین ابزاری که برای بررسی $Y=g(X)$ داشتیم LOTUS بود که فقط $E(Y)$ را می‌داد نه شکل توزیع Y را.

کاربردها:

مثال: تعداد افراد بر اساس سانتیمتر از توزیع $Normal(140, 20)$ تبعیت می‌کند.
تعداد افراد بر اساس اینچ از چه توزیعی تبعیت می‌کند؟

مثال: نمرات خام دانشجویان از توزیع $Uniform(9, 20)$ تبعیت می‌کند. اگر
نمودار $Y=\sqrt{20X}$ را روی نمرات اعمال کنیم نمرات جدید از چه توزیعی تبعیت می‌کنند؟

مثال: اگر $X\sim Normal(0,1)$ تبعیت کند $Y=X^2$ را متغیر تصادفی با توزیع
مربع کای Chi-Squared می‌نامیم. تابع چگالی Y چیست؟
χ^2

مثال: اگر $X\sim Normal(\mu,\sigma^2)$ تبعیت کند و $Y=e^X$ آن گاه گوییم Y از
توزیع Log Normal تبعیت می‌کند. تابع چگالی احتمال Y چیست؟

قضیه ۱) اگر g تابع یک به یک (معکوس پذیر) و مشتق پذیر اکیداً صعودی یا اکیداً نزولی باشد آنگاه تابع PDF روی متغیر $Y=g(X)$ برابر است با:

$$f_Y(y) = f_X(x)\left|\frac{dx}{dy}\right| \qquad x = g^{-1}(y) \text{ که } y = g(x)$$

اثبات: فرض کنیم g اکیداً صعودی باشد. آنگاه:

$$F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = P(g^{-1}(g(X)) \leq g^{-1}(y))$$

$$= P(X \leq x) = F_X(x) = F_X(g^{-1}(y))$$

$$f_Y(y) = \frac{dF_Y(y)}{dy} = \frac{dF_X(g^{-1}(y))}{dy} = \frac{dF_X(x)}{dy} = f_X(x)\frac{dx}{dy}$$

طبق قانون زنجیره

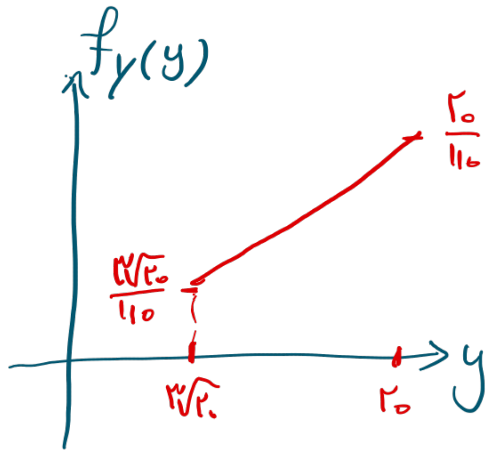
درحالتی که g تابع اکیداً نزولی باشد می‌توانیم نشان دهیم $f_Y(y) = -f_X(x)\frac{dx}{dy}$ که چون $\frac{dx}{dy}$ منفی است، می‌توانیم در حالت کلی (هم g اکیداً صعودی و هم اکیداً نزولی) از قدرمطلق استفاده کنیم:

$$f_Y(y) = f_X(x)\left|\frac{dx}{dy}\right|$$

مثال: اگر $Y = \sqrt{20X}$ و $X \sim Uniform(9, 20)$:

$$f_Y(y) = \frac{1}{20-9}\left|\frac{dx}{dy}\right| \qquad x = \frac{y^2}{20} \quad (y \geq 0) \Leftarrow y^2 = 20x$$

$$\frac{dx}{dy} = \frac{2y}{20} = \frac{y}{10}$$



$$\Rightarrow f_Y(y) = \frac{1}{11} \times \left|\frac{y}{10}\right| = \left|\frac{y}{110}\right| = \frac{y}{110}$$

فقط روی دامنه y برقرار است،

$$9 \le x \le 20 \to 3\sqrt{20} \le y \le 20$$

مثال: تابع چگالی احتمال مربع کای؟

$X \sim Normal(0,1)$

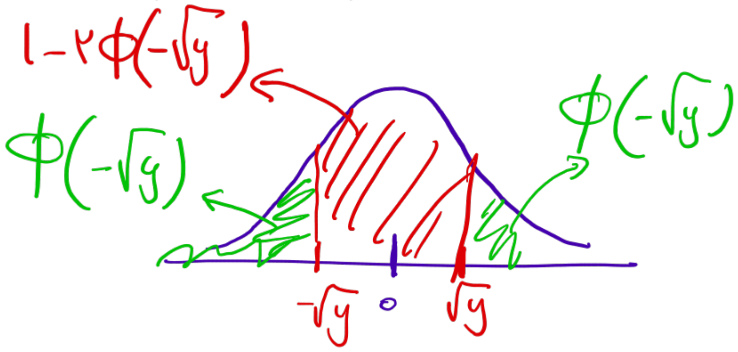
$Y = X^2 \longrightarrow$ تابع $g(x)=x^2$ معکوس‌پذیر نیست

اکیداً صعودی/نزولی هم نیست.

$y \ge 0$

$$F_Y(y) = P(Y \le y) = P(X^2 \le y) = P(-\sqrt{y} \le X \le \sqrt{y})$$

$$= 1 - 2\Phi(-\sqrt{y})$$



$$\Rightarrow f_Y(y) = \frac{d}{dy}\left(1 - 2\Phi(-\sqrt{y})\right) = -2\left(-\frac{1}{2\sqrt{y}}\right)\varphi(-\sqrt{y})$$

$$= \frac{1}{\sqrt{y}}\varphi(-\sqrt{y}) = \frac{1}{\sqrt{y2\pi}} e^{-\frac{y}{2}}$$