The classical one.

Probability

(90) The random variable X has an exponential distribution with parameter λ, so that

$$F_X(x) = P(X \le x) = 1 - e^{-\lambda x}.$$

Let $Y = e^{-\lambda X}$.

C. 17. Which option is equivalent to the event $Y \le y$.

A. $X \geq -\log y$ B. $X \le \log y$ $\boxed{\text{C. } X \geq -\lambda^{-1} \log y}$ D $X \geq \log(y/\lambda)$

E $X \geq 1 - e^{-\lambda y}$

18. Which option is the probability density function of Y for $0 \le y \le 1$?

A A $\boxed{\lambda^{-1} y^{\lambda^{-1}-1}}$ B $\lambda y^{-\lambda-1}$ C. $\lambda e^{-\lambda y}$ D $\lambda^{-1} y^{-\lambda^{-1}-1}$ C. $\lambda y^{\lambda-1}$

some transformation.

$$Y \le y \Leftrightarrow e^{-\lambda X} \le y \Leftrightarrow -\lambda X \le \log y \Leftrightarrow X \geq -\lambda^{-1} \log y.$$

① get the CDF

② get the PDF by derivatives.

For $0 \le y \le 1$. $P(Y \le y) = P(X \geq -\lambda^{-1} \log y) = 1 - F_X(-\lambda^{-1} \log y)$

$$= y^{\lambda^{-1}}.$$

Thus for $0 \le y \le 1$, $f_Y(y) = F_Y'(y) = \lambda^{-1} y^{\lambda^{-1}-1}$.

λ^{-1} 这个整体就是一个常数，无需做过多运算.

Summary

$\boxed{\text{Exponential distribution}}$ memoryless.

$$\frac{P\{X > s+t,\ X > t\}}{P\{X > t\}} = P\{X > s\}$$

$$F_X(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases} \qquad f_X(x) = \frac{dF(x)}{dx} = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

$$E(X) = \int_0^\infty x \lambda e^{-\lambda x} dx = \frac{1}{\lambda} \qquad Var(X) = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

$$E(X^2) = \int_0^\infty x^2 \lambda e^{-\lambda x} dx = \frac{2}{\lambda^2}$$

probability.

(9) Let X be an exponential random variable with parameter PDF

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

(a) If $t \in R$, find $P(X > t)$

(b) Prove that.

$P(X > s+t \mid X > s) = P(X > t)$ for all $s > 0$ and $t > 0$.

(c) Find the mode and median of X.

$F_X(x) = \int_{-\infty}^{x} f(x)dx$ ⇐ $x \geq 0$.

$F_X(x) = \int_0^x f(x)\,dx$ ⇐

$= \int_0^x \lambda e^{-\lambda x}\,dx$

$= [-e^{-\lambda x}]_0^x$

$= -e^{-\lambda x} + 1$

$= 1 - e^{-\lambda x}$

(a) ∵ $\boxed{t \in R}$ classify t.

$t < 0$. $P(X > t) = 1$

$t \geq 0$. $P(X \leq t) = 1 - e^{-\lambda t}$

$\boxed{P(X > t)} = 1 - P(X \leq t) = \boxed{e^{-\lambda t}}$

(b) $P(X > s+t \mid X > s) = \dfrac{P(X > s+t)}{P(X > s)} = \dfrac{e^{-\lambda(s+t)}}{e^{-\lambda s}} = e^{-\lambda t} = P(X > t)$ q.e.d.

(c). The mode y if $f_X(y) \geq f_X(x)$. for all $x \in R$.

since. $f_X(0) = 1$ and $f(x)$ is decreasing on $(0, \infty)$

so the mode is 0.

Median: $\boxed{P(X > m) = 0.5}$ $e^{-\lambda m} = 0.5$. $m = \lambda^{-1} \log 2$.

这个题最重要的就是要给t的值分个类.

Summary

$\boxed{F_X(x) = \int_{-\infty}^{x} f(x)\,dx.}$

sometimes need to specify it.

$\boxed{P(X > m) = 0.5}$ ⇒ MEDIAN

Pababrility.

(9a) A random variable X having a Beta distr. with parameter $\alpha>0$ and $\beta>0$ has probability density function.

$$f(x)=\begin{cases}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}, & 0<x<1\\ 0 & \text{otherwise..}\end{cases}$$

where for $t>0$

$$\Gamma(t)=\int_0^\infty y^{t-1}e^{-y}dy \text{ and } \Gamma(t+1)=t\Gamma(t).$$

(i) Prove that the mean and variance of X are $\frac{\alpha}{\alpha+\beta}$ and $\frac{\alpha\beta}{(\alpha+\beta)^\alpha(\alpha+\beta+1)}$, respectively.

(Hint). Note that $\boxed{\int_0^1 f(x)dx=1}$ for any $\alpha,\beta>0$. Rearranging this gives you an expression for $\int_0^1 x^{a-1}(1-x)^{b-1}dx$).

(ii) If $\alpha>1$ and $\beta>1$, find the mode of X.

$\Rightarrow$ (Hint): $\int_0^1\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}=1$

bring out the constant.

$$\int_0^1 x^{\alpha-1}(1-x)^{\beta-1}=\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

$$E(x)=\int_0^1 xf(x)dx=\int_0^1 x\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}dx.$$

$$=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\int_0^1 x^{(\alpha+1)-1}(1-x)^{\beta-1}dx.$$

$$=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\cdot\frac{\Gamma(\alpha+1)\Gamma(\beta)}{\Gamma(\alpha+\beta+1)}$$

$$=\frac{\alpha}{\alpha+\beta}.$$

$\boxed{\text{Since } \Gamma(t+1)=t\Gamma(t) \quad \Gamma(t)=(t-1)!}$

Summary

5

$$E(x^2) = \int_0^1 x^2 f(x)\,dx$$
$$= \int_0^1 x^2 \cdot \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot x^{\alpha-1}(1-x)^{\beta-1}\,dx$$
$$= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot \int_0^1 x^{(\alpha+2)-1}(1-x)^{\beta-1}\,dx$$
$$= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot \frac{\Gamma(\alpha+2)\Gamma(\beta)}{\Gamma(\alpha+\beta+2)} \qquad \Gamma(t) = (t-1)!$$
$$= \frac{\alpha(\alpha+1)}{(\alpha+\beta+1)(\alpha+\beta)}$$

$$Var(x) = E(x^2) - [E(x)]^2 = \frac{\alpha(\alpha+1)}{(\alpha+\beta+1)(\alpha+\beta)} - \left(\frac{\alpha}{\alpha+\beta}\right)^2 = \frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}$$

(ii): The mode must satify $\frac{dy}{dx} = 0$.

$\frac{df}{dx} = 0$. (product rule.)

$$\left[(\alpha-1)x^{\alpha-2}(1-x)^{\beta-1} + x^{\alpha-1}(1-x)^{\beta-2}(\beta-1)\right]\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} = 0.$$

don't forget the "−" it's actually a composition function.

$$(\alpha-1)x^{\alpha-2}(1-x)^{\beta-1} - x^{\alpha-1}(1-x)^{\beta-2}(\beta-1) = 0$$
$$(\alpha-1)(1-x)^{\beta-1} - x(1-x)^{\beta-2}(\beta-1) = 0.$$
$$(\alpha-1)(1-x) - x(\beta-1) = 0.$$
$$x = \frac{\alpha-1}{\alpha+\beta-2} = (\alpha-1)/\alpha+\beta-2$$

Summary

哈哈我刚刚做的时候好像捣腾了一个新方法.

$$P_X(s) = E[s^k] = \sum_{k=1}^{\infty} s^k (1-p)^{k-1} p$$

$$= sp \sum_{k=1}^{\infty} [s(1-p)]^{k-1}$$

$$= \frac{sp}{[1-s(1-p)]} \sum_{k=1}^{\infty} [s(1-p)]^{k-1} [1-s(1-p)]$$

$$= \frac{sp}{1-s(1-p)}.$$

我直接把后面配平了 哈哈哈....

Probability generating function (Example)

If $X \sim Geo(p)$.

$$P_X(s) = E[s^X] = \sum_{k=1}^{\infty} s^k \cdot (1-p)^{k-1} p.$$

$$= p \sum_{k=1}^{\infty} s^k (1-p)^{k-1}$$

$$= \frac{p}{1-p} \sum_{k=1}^{\infty} s^k (1-p)^k$$

$$= \frac{p}{1-p} \sum_{k=1}^{\infty} [s(1-p)]^k$$

$$\sum_{k=1}^{\infty} q^k = \frac{q}{1-q} \quad (-1<q<1 \rightarrow \text{converge}).$$

$$\sum_{k=1}^{n} q^k = q^1 + q^2 + \cdots + q^n = \frac{q - q^{n+1}}{1-q}$$

$$\lim_{n\to\infty} \sum_{k=1}^{n} q^k = \frac{q}{1-q}.$$

$$= \frac{p}{1-p} \cdot \frac{s(1-p)}{1-s(1-p)}$$

$$= \frac{sp}{1-s(1-p)}. \quad (-1 < s(1-p) < 1)$$

Let $Y \sim NB(n, p)$. (Negative Binomial)

trials until n^{th} success.

for every trial in NB, it is a Geometric dist.

$Y = X_1 + \cdots + X_n$. $X_i \sim Geo(p)$.

$$\cancel{\sum_{k=n}^{\infty} \binom{k-1}{n-1} (1-p)^{k-n} p^n \quad (-1<s<1)}$$

$$\cancel{= \sum_{k=n}^{\infty}}$$

$$\cancel{P_Y(s) = E[s^Y] = \sum_{k=n}^{\infty} s^k \binom{k-1}{n-1} (1-p)^{k-n} p^n \quad (-1<s<1)}$$

$$\cancel{= (sp)^n \sum_{k=n}^{\infty} \binom{k-1}{n-1} (1-p)^{k-n}}$$

Summary

$\boxed{\text{Negative Binomial distribution.}}$

$$P_X(k) = \binom{k-1}{r-1} (1-p)^{k-r} p^r. \quad k = r, r+1, \ldots$$

$$E(X) = \frac{r}{p}$$

$$var(X) = \frac{r(1-p)}{p^2}$$

$$P_Y(s) = E[s^Y] = \sum_{k=n}^{\infty} s^k \binom{k-1}{n-1} (1-p)^{k-n} p^n \quad (-1 < s < 1)$$

① $$\sum_{k=n}^{\infty} \binom{k-1}{n-1} [s(1-p)]^{k-n} (sp)^n \rightarrow \text{sum is not } 1$$

$$= \frac{(sp)^n}{[1-s(1-p)]^n} \sum_{k=n}^{\infty} \underbrace{\binom{k-1}{n-1} [s(1-p)]^{k-n} [1-s(1-p)]^n}_{\text{pmf of NB}}.$$

$$= \frac{(sp)^n}{[1-s(1-p)]^n}$$

$0 < s(1-p) < 1$

$0 < s < \frac{1}{1-p}$

This method only show $0 < s < 1$, but can never show $-1 < s < 0$.

② $Y = X_1 + \cdots + X_n \quad X_i \sim Geo(p)$

$$P_X = \frac{ps}{1-s+ps}$$

$$P_Y(s) = \prod_{i=1}^{n} P_{X_i}(s) = \frac{(ps)^n}{(1-s+ps)^n} \quad (-1 < s < 1)$$

生成函数中一个性质的运用.

我发现两个推导过程都是用 s^k 去妥协 $(1-p)$ 这个变量，相当于依照原函数的样子把 s 当一个新变量靠在一起，最后再找一个和新变量相加为 1 的变量凑成新的 pmf.

Summary

Lemma.

$Z = X_1 + X_2 + \cdots + X_n.$

$$\boxed{P_Z(s) = \prod_{i=1}^{n} P_{X_i}(s)}$$

5 . .

Probability.

continuous.

How to integrate.

(9) Let X and Y be continuous random varibles with joint distribution function.

$F(x,y) = kxy(x+y) \quad 0<x<1, \; 0<y<1$

(a) Evaluate the constant K, and find the joint pdf of X and Y.

(b) Find the marginal and conditional pdfs of X

(c) Evaluate $P(X<0.5, Y<0.5)$ $P(X<0.5)$ $P(Y<0.5)$.

$\because F(\infty,\infty)=1 \quad F(\infty,\infty)=F(1,1)=1$

$\therefore 2K=1 \quad K=\frac{1}{2}$.

get the pdf. derivative.

$\boxed{f(x,y)} = \frac{\partial^2}{\partial x \partial y} F(x,y) = \frac{\partial^2}{\partial x \partial y}\left\{\frac{x^2 y}{2} + \frac{xy^2}{2}\right\}$

$$= \begin{cases} x+y & 0<x,y<1 \\ 0 & \text{otherwise.} \end{cases}$$

pdf of X. ⇒ integrate pdf.

$\therefore \boxed{f_X(x)} = \int_{-\infty}^{\infty} f(x,y)\,dy = \int_0^1 x+y \, dy$

$$= \begin{cases} x+\frac{1}{2} & \text{if } 0<x<1 \\ 0 & \text{otherwise} \end{cases}$$

$\boxed{f_Y(y)} = \int_{-\infty}^{\infty} f(x,y)\,dx = \int_0^1 x+y \, dx = \begin{cases} y+\frac{1}{2}, & 0<y<1 \\ 0, & \text{otherwise.} \end{cases}$

求具体条件概率. 先求pdf ←

$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)} = \begin{cases} \frac{x+y}{\frac{1}{2}+y} & \text{if } 0<x<1 \\ 0 & \text{otherwise.} \end{cases}$

(c) $P(X<0.5, Y<0.5) = F(\frac{1}{2},\frac{1}{2}) = \frac{1}{8}$

$P(X<0.5) = \int_0^{\frac{1}{2}} f_X(x)\,dx = \int_0^{\frac{1}{2}} \frac{1}{2}+x \, dx = \frac{3}{8}$

$P(Y<0.5) = \int_0^{\frac{1}{2}} f_Y(y)\,dy = \frac{3}{8}$.

Summary

Probability.

(94) The joint probability density function of X and Y is given by.

$$f_{X,Y}(x,y) = \frac{6}{7}\left(x^2 + \frac{xy}{2}\right) \quad 0<x<1,\ 0<y<2$$

Find $P(Y>1 \mid X<\frac{1}{2})$

$$P(Y>\tfrac{1}{2},\ X<\tfrac{1}{2}) = \int_{x=0}^{\frac{1}{2}}\int_{y=\frac{1}{2}}^{2} \frac{6}{7}\left(x^2+\frac{xy}{2}\right) dy\, dx$$

$$= \frac{69}{448}$$

$$P(X<\tfrac{1}{2}) = \int_{x=0}^{\frac{1}{2}}\int_{y=0}^{2} \frac{6}{7}\left(x^2+\frac{xy}{2}\right) dy\, dx$$

$$= \frac{5}{7\times 2^2}$$

$$P(Y>\tfrac{1}{2} \mid X<\tfrac{1}{2}) = \frac{P(Y>\frac{1}{2},\ X<\frac{1}{2})}{P(X<\frac{1}{2})} = \frac{69}{80}$$

Probability

(85) $$F_Y(y) = \begin{cases} 0 & y \le 1 \\ (y-1)^4 & 1<y\le 2 \\ 1 & y>2 \end{cases}$$

Find $P(\frac{3}{4}<Y<\frac{7}{4})$

defective: $F(\frac{7}{4}) - \boxed{F(\frac{3}{4})} = \frac{5}{16}$ → it's not in the range!!!!!!!

correct: $F(\frac{7}{4}) - 0 = (\frac{7}{4}-1)^4 - 0 = \frac{81}{256}$

这个题太容易又坑了. 保险的做法就是求出了 $F_Y(y)$. 然后. 代值过程一个个检验

it's easy to just use the function directly.

We MUST CHECK WHETHER IT IS IN THIS RANGE

Summary

5

Probability.

先计算 pdf
再积分

(98) $f_{X,Y}(x,y) = \begin{cases} \frac{1}{4}x^2y^2(4y+x) & 0\le x\le 2,\ 0\le y\le 1 \\ 0 & \text{otherwise.} \end{cases}$

Find. $P(X\geq 1 \mid Y=\frac{1}{2})$

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

$$f_{X|Y}(x|\tfrac{1}{2}) = \frac{f_{X,Y}(x,\frac{1}{2})}{f_Y(\frac{1}{2})}$$

$$f_{X,Y}(x,\tfrac{1}{2}) = \frac{1}{8}x^2 + \frac{1}{16}x^3$$ 直接代入.

$$f_Y(y) = \int_0^2 f(x,y)\,dx = \frac{7}{12}$$

$$f_{X|Y}(x|\tfrac{1}{2}) = \frac{f_{X,Y}(x,\frac{1}{2})}{f_Y(\frac{1}{2})} = \frac{\frac{1}{8}x^2+\frac{1}{16}x^3}{\frac{7}{12}} = \begin{cases}\frac{3}{28}(2x^2+x^3) & 0\le x\le 2\\ 0 & \text{otherwise.}\end{cases}$$

$$P(X>1 \mid Y=\tfrac{1}{2}) = \int_1^{\infty} f_{X|Y}(x|1/2)\,dx = \int_1^2 \frac{3}{28}(2x^2+x^3)\,dx = \frac{101}{112}$$

Probability.

(99) Three card: One is red on both sides / One is black on both sides / One is red, one is black. Card is shuffled and one is chosen at random. If the upper side is red, what is the probability that the other side is red?

key step. Let the card be $\boxed{\text{event}}$ ⇐

$\boxed{\text{Let } F = \text{'Upper is red'} \quad E_1 = RR \quad E_2 = BB \quad E_3 = RB \text{ card.}}$

We want $P(E_1|F)$ $\quad P(E_1)=P(E_2)=P(E_3) = 1/3$

$\boxed{P(F|E_1) = 1 \quad P(F|E_2) = 0 \quad P(F|E_3) = \frac{1}{2}}$

$$P(E_1/F) = \frac{P(F|E_1)P(E_1)}{P(F|E_1)P(E_1)+P(F|E_2)P(E_2)+P(F|E_3)P(E_3)} = \frac{2}{3}$$

现在还是规范点写吧.
不要用 RR RB BB 了.

Summary

用脚标代替吧 E_1 E_2 E_3.

概率论易错知识点

A probability measure is a real-valued set function P defined on the events of a sample space Ω which satisfies the following three axioms.

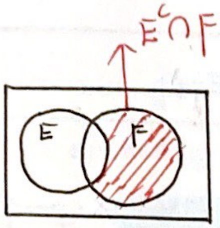
Three axioms.

A_1: $P(E) \geq 0$ for any event E

A_2: $P(\Omega) = 1$

A_3: If $E_1, E_2 \ldots E_n$ are s.t. $E_i \cap E_j = \phi$ for $i \neq j$. then

$$P(E_1 \cup E_2 \cup \cdots) = P(E_1) + P(E_2) + \cdots\cdots$$



Common trick: use to prove $P(E \cup F) = P(E) + P(F) - P(E \cap F)$

$$E \cup F = E \sqcup (F \cap E^c) \qquad F = (F \cap E) \sqcup (F \cap E^c)$$

我可以先写F. 之后E的目的不过是试图去减去$(F \cap E^c)$.

ordered arrangement ←

• Permutation of length r of n objects

$$^nP_r = n(n-1)(n-2)\cdots(n-r+1) = \frac{n!}{(n-r)!} \qquad ^nP_n = n!$$

unordered arrangement. ←

• Combination of r objects from a collection of n objects.

$$^nC_r = \binom{n}{r} = \frac{^nP_r}{r!} = \frac{n!}{(n-r)!\,r!} \qquad \binom{n}{r} = \binom{n}{n-r}$$

像这种如此奇怪的分母 我们会想方设法把他消掉

$$\binom{n}{m} = \binom{n-1}{m-1} + \binom{n-1}{m} \qquad \binom{n}{m} = \frac{n}{m}\binom{n-1}{m-1} \qquad \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{m} = 2^n$$

$$P(A^c|F) = 1 + \frac{P(A^c \cap F)}{P(F)} - 1$$
$$= 1 + \frac{P(A^c \cap F) - P(F)}{P(F)}$$
$$= 1 - \frac{P(F) - P(A^c \cap F)}{P(F)}$$
$$= 1 - \frac{P(A \cap F)}{P(F)}$$
$$= 1 - P(A|F)$$

$$P(\Omega|F) = 1 \qquad P(\phi|F) = 0$$

$P(A^c|F) = 1 - P(A|F)$ proof by common trick.

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

$$P(A \cap B \cap C) = P(C|A \cap B)\underbrace{P(B|A)P(A)}_{P(AB)}$$

$P(ABC)$

可以倒着顺序先写 $P(A)$.

portition can be finite or infinite. ⇐

Theorem of Total Probability ⇒ 是一组事件，说明划分是针对Ω而言.

Definition: A partition of a sample space Ω is a collection of events $E_1, E_2, \cdots$ in Ω s.t.

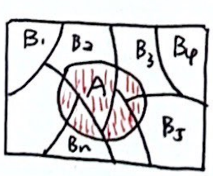
(i) $E_i \cap E_j = \phi$ for all $i \neq j$ (disjoint set)

(ii) $\bigsqcup_i E_i \; (= E_1 \sqcup E_2 \sqcup \cdots) = \Omega$

Summary

Let $F \subseteq \Omega$

$$\boxed{P(F) = \sum_i P(F|E_i)P(E_i)}$$



Baye's Theorem.

Simple form: $P(E|F) = \frac{P(F|E)P(E)}{P(F)}$ $P(E)>0$ and $P(F)>0$

General form: if E_1, E_2 .. form a partition of Ω

$$\boxed{P(E_i|F) = \frac{P(F|E_i)P(E_i)}{\sum_i P(F|E_i)P(E_i)}}$$

Mutually independent: $P(\bigcap_j E_j) = \prod_j P(E_j)$

Probability mass function of a random variable X (pmf)

$$\boxed{P_X(x) = P(X=x) = P(\{w: X(w)=x\})}$$

$E_i = \{w: X(w)=x_i\}$. Then $\{E_i\}$ is a partition of Ω. so

$$\boxed{\sum_i p_X(x_i) = \sum_i P(E_i) = P(\bigcup_i E_i) = P(\Omega) = 1}$$

Cumulative distribution function (c.d.f.)

$$\boxed{F_X(x) = F(x) = P(X \le x) = \sum_{y \le x} P_X(y) \quad x \in R}$$

(i) F_X is non-decreasing, i.e. if $x \le y$. $F_X(x) \le F_X(y)$

(ii) $P_X(x_i) = F_X(x_i) - F_X(x_{i-1})$;

(iii) if $a<b$ $P(a<X\le b) = P(X\le b) - P(X \le a)$

本身是不会被减掉的 ← $= F_X(b) - F_X(a)$

一定要先明确是不是离散的变量.

(iv) $P(a \le X < b) = P(X<b) - P(X \le a) = \sum_{y \in [a,b)} P_X(y)$

(v) $P(a \le X < b) = P(X=a) + \sum_{y \in (a,b)} P_X(y)$

(vi) $F_X(\infty) = P(X \le \infty) = 1$ $F_X(-\infty) = P(X \le -\infty) = 0$

Mode: the mode of a discrete random variable is the outcome with the highest probability.

Summary

5 . .

The variance gives a measure of how $\boxed{\text{spread out}}$ the mass function of X is.

※: For r = 1, 2, 3. the r^{th} moment of X is defined $E[X^r]$.

Expectation : $\boxed{E(X) = \sum_x x\, p(x)}$ (also called the mean)

Lemma : g(x) is a Function of X. $\boxed{E[g(x)] = \sum_x g(x) p_X(x)}$

Properties: ① if g(x) = b. a constant. $E[g(x)] = b$

② For all a. b ∈ R. $E[ag(x) + bh(x)] = aE[g(x)] + bE[h(x)]$

③ $E[aX+b] = aE(X) + b$

Variance : $Var(X) = E[(X - E(X))^2] = \sum_x [x - E(X)]^2 p_X(x)$.

Standard deviation of X is $\sqrt{Var(X)}$

Properties: ① $Var(X) = E(X^2) - [E(X)]^2$

② $Var(X) \geq 0$ $\quad Var(X) = 0 \iff P(X = E(X)) = 1$

③ For all a, b ∈ R $\quad Var(aX+b) = a^2 Var(X)$

Further properties

① $E(X+Y) = E(X) + E(Y)$

② $E(X_1 + X_2 + \cdots + X_n) = E(X_1) + E(X_2) + \cdots + E(X_n)$

③ if X, Y are independent then $Var(X+Y) = Var(X) + Var(Y)$.

④ if $X_1, X_2 \cdots X_n$ are mutually independent, then

$Var(X_1 + X_2 + \cdots + X_n) = Var(X_1) + Var(X_2) + \cdots + Var(X_n)$.

Summary

STANDARD DISCRETE DISTRIBUTION

Bernoulli Distribution: A Bernoulli trial is a simple random experiment with two outcomes: success or failures

$$X = \begin{cases} 1 & \text{success} \quad p \\ 0 & \text{failure} \quad 1-p \end{cases}$$

$E(X) = p \times 1 + 0 \times (1-p) = p$

$Var(X) = p(1-p)$

Binomial Distribution: n independent Bernoulli trials

$X \sim Bin(n, p)$

$$\boxed{P_X(k) = P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}} \quad k = 0, 1, \cdots n$$

$$\boxed{E(X) = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k} = np}$$

$$\boxed{var(X) = np(1-p)}$$

$F_X(x) = P(X \le x) = \boxed{\sum_{k=0}^{[x]}} \binom{n}{k} p^k (1-p)^{n-k}$ (use table)

If $X \sim Bin(n, p)$, $Y = n - X$. $Y \sim Bin(n, 1-p)$.

$P_Y(k) = P_X(n-k) = \binom{n}{n-k} p^{n-k} (1-p)^{n-(n-k)}$

$= \binom{n}{k} (1-p)^k p^{n-k}$.

Proof of $E(X)$:

$E(X) = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k}$

$= \sum_{k=0}^{n} k \cdot \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}$

$= \sum_{k=0}^{n} \frac{n!}{(k-1)!(n-k)!} p^k (1-p)^{n-k}$

$= \sum_{k=0}^{n} n \frac{(n-1)!}{(k-1)![(n-1)-(k-1)]!} p^k (1-p)^{n-k}$

$= \sum_{k=0}^{n} np \frac{(n-1)!}{(k-1)![(n-1)-(k-1)]!} p^{k-1} (1-p)^{(n-1)-(k-1)}$

$= np \sum \text{-----}$

$= np$.

$E[X(X-1)]$

$= \sum_{k=0}^{n} k(k-1) \binom{n}{k} \cdot p^k (1-p)^{n-k}$

$= \sum_{k=0}^{n} k(k-1) \frac{n!}{(n-k)!\,k!} \cdot p^k (1-p)^{n-k}$

$= \sum_{k=0}^{n} p^2 n(n-1) \frac{(n-2)!}{(k-2)![(n-2)-(k-2)]!} p^{k-2} (1-p)^{[(n-2)-(k-2)]}$

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

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

$= \boxed{E[X(X-1)]} + E(X) - [E(X)]^2$

$= n(n-1)p^2 + np - n^2p^2$

$= np(1-p)$.

Summary

这样做的好处就是自动把它和阶乘联系在了一起.

只是这种有(n)的. 用 $E(X(X-1))$ 可以直接消掉阶乘.

Simpler Method ①

Proof: $E(X)$.

$E(X)=\sum_{k=1}^{\infty} k\cdot(1-p)^{k-1}p$

$=\sum_{k=1}^{\infty} k\cdot q^{k-1}p$, $q=1-p$

$=\sum_{k=1}^{\infty}(k-1+1)q^{k-1}p$

$=\sum_{k=1}^{\infty}(k-1)q^{k-1}p+\underbrace{\sum_{k=1}^{\infty}q^{k-1}p}_{1}$

$=\sum_{j=0}^{\infty} j q^{j}p+1$

$=\boxed{q}\sum_{j=1}^{\infty} j q^{j-1}p+1$

$=qE(X)+1$

$E(X)[1-q]=1$ $E(X)=\frac{1}{p}$

$E(X^2)=\sum_{k=1}^{\infty}k^2q^{k-1}p$

$=\sum_{k=1}^{\infty}(k-1+1)^2q^{k-1}p$

$=\sum_{k=1}^{\infty}[(k-1)^2+2(k-1)+1]q^{k-1}p$

$=\sum_{k=1}^{\infty}(k-1)^2q^{k-1}p+2\sum_{k=1}^{\infty}(k-1)q^{k-1}p+\sum_{k=1}^{\infty}q^{k-1}p$

$=q\sum_{j=1}^{\infty}j^2q^{j-1}p+2q\sum_{j=1}^{\infty}jq^{j-1}p+1$

$=\boxed{q}E(X^2)+\boxed{2q}E(X)+1$

$E(X^2)=\frac{2q+p}{p^2}$ 提出q是关键

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

$=\frac{1-p}{p^2}$

Method ②

差比数列.作差.求极限.

Method ③

$E(X)=\sum_{k=1}^{\infty}k(1-p)^{k-1}p$ 求积分后再求导

$=p\sum_{k=1}^{\infty}\frac{d}{dp}\{-(1-p)^k\}$

$=-p\frac{d}{dp}\sum_{k=1}^{\infty}(1-p)^k$ 等比数列 $\sum_{k=1}^{\infty}q^k$. $\sum_{k=1}^{\infty}\frac{q-q^{k+1}}{1-q}=\frac{q}{1-q}$

$=-p\frac{d}{dp}\left(\frac{1-p}{p}\right)$

$=-p\cdot\frac{-1}{p^2}$

$=\frac{1}{p}$

Geometric distribution.

the number of trials until the first success appear.

$X\sim Geom(p)$.

$$\boxed{p_X(k)=(1-p)^{k-1}p,\quad k=1,2,3,\dots}$$

(k^{th} trial is the first success then the first $(k-1)$ trials must have been failures)

$$\boxed{\begin{aligned}&E(X)=\frac{1}{p}\\&var(X)=E[X(X-1)]+E(X)-[E(X)]^2=\frac{1-p}{p^2}\end{aligned}}$$

$E[X(X-1)]=\frac{2(1-p)}{p^2}$,

Poisson Distribution

$$\boxed{p_X(k)=e^{-\lambda}\frac{\lambda^k}{k!}},\quad k=\boxed{0},1,2,\dots$$

$X\sim Po(\lambda)$. $\sum_{k=0}^{\infty}p_X(k)=e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}=e^{-\lambda}e^{\lambda}=1$

$$\boxed{\begin{aligned}&E(X)=\sum_{k=0}^{\infty}k\cdot e^{-\lambda}\frac{\lambda^k}{k!}=\lambda e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!}=\lambda e^{-\lambda}\cdot e^{\lambda}=\lambda\\&E[X(X-1)]=\sum_{k=0}^{\infty}k(k-1)\cdot e^{-\lambda}\frac{\lambda^k}{k!}=\lambda^2e^{-\lambda}\sum_{k=2}^{\infty}\frac{\lambda^{k-2}}{(k-2)!}=\lambda^2e^{-\lambda}\cdot e^{\lambda}=\lambda^2\\&var(X)=E[X(X-1)]+E(X)-[E(X)]^2=\lambda.\end{aligned}}$$

$F_X(x)=\sum_{0\le k\le x}e^{-\lambda}\frac{\lambda^k}{k!}$. (use table).

Derivatives of Poisson Distribution

$\frac{\;}{n}$ p. $E(X)=np=\lambda$.

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

$p(X=k)=\binom{n}{k}p^k\cdot(1-p)^{n-k}$

$=\frac{n!}{(n-k)!\,k!}\frac{\lambda^k}{n^k}\cdot\left(1-\frac{\lambda}{n}\right)^{n-k}$

$=\frac{n\cdot(n-1)\cdots(n-k+1)}{k!}\cdot\frac{\lambda^k}{n^k}\cdot\frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^k}$

$=\frac{n(n-1)\cdots(n-k+1)}{n^k}\cdot\frac{\lambda^k}{k!}\cdot\frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^k}$

$=1\cdot\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)\cdot\frac{\lambda^k}{k!}\cdot\frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^k}$

$=1\cdot\frac{\lambda^k}{k!}\left(1-\frac{\lambda}{n}\right)^n=\frac{\lambda^k}{k!}e^{-\lambda}$

$$\begin{cases}\lim\left(1+\frac{x}{n}\right)^n=e^x\\ \sum_{k=0}^{\infty}\frac{a^k}{k!}=e^a\end{cases}$$

Summary

$E[X(X-1)]$

$=\sum_{k=1}^{\infty}k(k-1)(1-p)^{k-1}p$

$=p(1-p)\sum_{k=1}^{\infty}k(k-1)(1-p)^{k-2}$

$=p(1-p)\sum_{k=1}^{\infty}\frac{d^2}{dp^2}\{(1-p)^k\}$

$=p(1-p)\frac{d^2}{dp^2}\sum_{k=1}^{\infty}\{(1-p)^k\}$

$=p(1-p)\frac{d^2}{dp^2}\left\{\frac{1-p}{p}\right\}$

$=p(1-p)\cdot 2p^{-3}$

$=\frac{2p(1-p)}{p^3}$

$var(X)=E[X(X-1)]+E(X)-E^2(X)$

$=\frac{2p(1-p)}{p^3}+\frac{1}{p}-\frac{1}{p^2}$

$=\frac{1-p}{p^2}$

$k = 10$ times $r = 8$

$r-1$ 3 success.

× × × × × × × V V V must success.

$\binom{k-1}{r-1}$

$n \uparrow$: $p \downarrow = E(x)$.

Poisson approximation $\boxed{\text{large } n. \text{ small } p.}$

$Bin(n,p) \approx Po(np)$

Negative Binomial distribution: the number of trials until r success.

$$\boxed{P_X(k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}, \quad k = r, r+1, r+2 \cdots}$$

(must have k^{th} trial successful, and $r-1$ successes in first $k-1$ trials)

$X \sim NB(p) \Rightarrow X_i \sim Geom(p)$.

$E(X) = E[X_1 + X_2 + \cdots + X_r] = \frac{r}{p}$

$var(X) = var[X_1 + \cdots + X_n] = \frac{r(1-p)}{p^2}$

Discrete Uniform distribution

$$\boxed{P_x(x) = \frac{1}{n} \quad x = 1, 2, \cdots n}$$

$E(X) = \sum_1^n x p = (1+2+3+\cdots+n) \cdot \frac{1}{n} = \frac{n(n+1)}{2} \cdot \frac{1}{n} = \frac{n+1}{2}$

$E(X^2) = \sum_1^n x^2 p = (1^2+2^2+3^2+\cdots+n^2) \frac{1}{n} = \boxed{\frac{n(n+1)(2n+1)}{6}} \cdot \frac{1}{n} = \frac{(n+1)(2n+1)}{6}$

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

$= \frac{(n+1)(2n+1)}{6} - \left(\frac{n+1}{2}\right)^2 = \frac{n^2-1}{12}$

Hypergeometric distribution. $n \le A+B$.

$P_X(k) = \frac{\binom{A}{k}\binom{B}{n-k}}{\binom{A+B}{n}} \quad k = 0, 1, \cdots n$

$E(X) = \frac{An}{A+B} \quad var(X) = \frac{AB\,n(A+B-n)}{(A+B)^2(A+B-1)}$

Summary

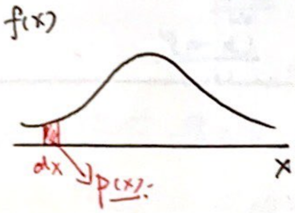
{ CONTINUOUS DISTRIBUTIONS }

Cumulative distribution function (cdf)

$$F_X(x) = P(X \le x) = \begin{cases} 0 & x < 0 \\ x & 0 \le x < 1 \\ 1 & x \ge 1 \end{cases}$$

Probability density function (pdf) of continuous rand. v.

$$\boxed{F_X(x) = \int_{-\infty}^{x} f_X(y)\, dy \quad \text{for all } x \in \mathbb{R}}$$



① $f_X(x) \ge 0$

② $\int_{-\infty}^{\infty} f_X(y)\, dy = F_X(\infty) = 1$

③ $\boxed{f_X(x) = \frac{dF_X(x)}{dx}}$

④ $f_X(x) = \frac{dF_X(x)}{dx} \approx \frac{F_X(x+h) - F_X(x)}{h}$

$\Rightarrow h f_X(x) \approx F_X(x+h) - F_X(x) = P(X \le x+h) - P(X \le x)$

$= P(x < X \le x+h)$

Expectation for continuous r.v / Variance

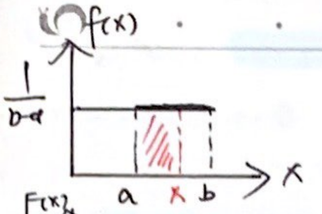
$$\boxed{\begin{aligned} &E(X) = \int_{-\infty}^{\infty} x f_X(x)\, dx \qquad E[g(x)] = \int_{-\infty}^{\infty} g(x) f_X(x)\, dx \\ &\text{var}(X) = E(X^2) - [E(X)]^2 \end{aligned}}$$

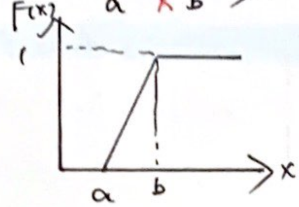
Median: is defined as x_0, s.t. $F_X(x_0) = \boxed{0.5}$

Check by $\frac{d}{dx} f(x) = 0$. ⇐ **Mode**: of X is point at which f_X is maximised i.e.

mode is x_0 if and only if $f_X(x_0) \ge f_X(x)$ for all x

Summary





STANDARD CONTINUOUS DISTRIBUTIONS

Uniform distribution. $X \sim U(a,b)$

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

$$F_X(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \le x < b \\ 1 & x \geq b \end{cases}$$

$E(X) = \frac{a+b}{2}$

$var(X) = \frac{(b-a)^2}{12}$

Exponential distribution

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

自然是趋近1.

$$F_X(x) = \begin{cases} 1 - e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

$E(X) = \int_0^\infty x \lambda e^{-\lambda x} dx = \frac{1}{\lambda}$ $E(X^2) = \int_0^\infty x^2 \lambda e^{-\lambda x} = \frac{2}{\lambda^2}$ (there).

$var(X) = E(X^2) - [E(X)]^2 = \frac{2}{\lambda^2} - (\frac{1}{\lambda})^2 = \frac{1}{\lambda^2}$

$E(X) = \int_0^\infty x \cdot \lambda e^{-\lambda x} dx$ By parts.

$= [-\lambda \cdot e^{-\lambda x}]_0^\infty - \frac{1}{\lambda}[e^{-\lambda x}]_0^\infty$

$= -\frac{1}{\lambda}[0-1]$

$= \frac{1}{\lambda}$.

$E(X^2) = \int_0^\infty x^2 \lambda e^{-\lambda x} dx$

$= [x^2 \cdot (-e^{-\lambda x})]_0^\infty + \int_0^\infty 2x \cdot e^{-\lambda x} dx$

$= 2\int_0^\infty x \cdot e^{-\lambda x} dx$

$= 2[[x \cdot (-\frac{1}{\lambda}) e^{-\lambda x}]_0^\infty + \int_0^\infty \frac{1}{\lambda} e^{-\lambda x} dx]$

$= \frac{2}{\lambda} \int_0^\infty e^{-\lambda x} dx$

$= \frac{2}{\lambda} \cdot [-\frac{1}{\lambda} e^{-\lambda x}]_0^\infty$

$= \frac{2}{\lambda} \cdot [0 + \frac{1}{\lambda}]$

$= \frac{2}{\lambda^2}$

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

$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}$, $x \in R$

$E(X) = \int_{-\infty}^{\infty} x f_X(x) dx = \mu$ 用换元法把 $x-\mu$ 换成 y.

$E(X^2) = \int_{-\infty}^{\infty} x^2 f_X(x) dx = \sigma^2 + \mu^2$ $y = x - \mu$.

$var(X) = \sigma^2 + \mu^2 - \mu^2 = \sigma^2$

我记得当时找了一个如何推导出 Normal Dis. 的视频，上面讲的我也没听太明白。只知道 $f(x) = K e^{\frac{1}{2} c x^2}$

Summary

c 为负，才收敛 $c = -\frac{1}{\sigma^2}$ $K = \frac{1}{\sqrt{2\pi}\sigma}$

$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{x^2}{2\sigma^2}}$ $x = x - \bar{x} = x - \mu$.

不过现在直接可以记住了.

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

$$E(X) = \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

Let $y = x-\mu$. $dy = dx$

$$E(X) = \int_{-\infty}^{\infty} (y+\mu) \cdot \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{y^2}{2\sigma^2}} \cdot dy$$

constant.

$$= \int_{-\infty}^{\infty} y \cdot \frac{1}{\sqrt{2\pi}\cdot\sigma} \cdot e^{-\frac{y^2}{2\sigma^2}} dy + \int_{-\infty}^{\infty} \mu \cdot \underline{\frac{1}{\sqrt{2\pi}\cdot\sigma} \cdot e^{-\frac{y^2}{2\sigma^2}} dy}$$

$= 0 + \mu = \mu$ pdf. $= 1$.

Let $y = \frac{x-\mu}{\sigma}$ $\frac{dy}{dx} = \frac{1}{\sigma}$

转换成 S.N.

$$E(X^2) = \int_{-\infty}^{\infty} (\sigma y+\mu)^2 \cdot \frac{1}{\sqrt{2\pi}\cdot\sigma} \cdot e^{-\frac{y^2}{2}} \cdot \sigma\, dy$$

$$= \int_{-\infty}^{\infty} (\sigma^2 y^2 + 2\sigma y\mu + \mu^2) \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{y^2}{2}} dy$$

$$= \int_{-\infty}^{\infty} \sigma^2 y^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy + \int 2\sigma\mu y \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{y^2}{2}} \cdot dy + \int \mu^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy$$

$= \sigma^2 + 2\sigma\mu \cdot 0 + \mu^2 = \sigma^2 + \mu^2$

$\mathrm{var}(X) = E(X^2) - [E(X)]^2 = \sigma^2 + \mu^2 - \mu^2 = \sigma^2$

Standard normal distribution

$f_X(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}$, $x \in R$ $\mu = 0$. $\sigma = 1$

$F_X(x) = \Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy$. (table).

- $P(a \le Z \le b) = \Phi(b) - \Phi(a)$
- $P(Z \ge a) = P(Z > a) = 1 - \Phi(a)$

Proof: Note $\boxed{X = \sigma Z + \mu}$

Z is standard normal

$Y = a(\sigma Z + \mu) = (a\sigma)Z + (a\mu + b)$

$Y \sim N(a\mu + b, a^2\sigma^2)$

↓

考虑如何用Z将X表示出来，Z是一个已知量

Further transformation $X \sim N(\mu, \sigma^2)$ $Y = aX + b$

$Y \sim N(a\mu + b, a^2\sigma^2)$

More Generally. $X_1 \cdots X_n$ be indep. r.v's

$Y = a_1 X_1 + a_2 X_2 + \cdots + a_n X_n$ $Y \sim N(\mu, \sigma^2)$

$\mu = a_1\mu_1 + \cdots + a_n\mu_n$

$\sigma^2 = a_1^2\sigma_1^2 + \cdots + a_n^2\sigma_n^2$

Summary

$$E(x) = \int_0^\infty \frac{1}{\Gamma(\alpha)} \beta^\alpha \cancel{x}^{(\alpha+1)-1} e^{-\beta x} dx$$

$$= \frac{1}{\Gamma(\alpha)} \int_0^\infty \beta^\alpha x^{(\alpha+1)-1} e^{-\beta x} dx$$

$$= \frac{\Gamma(\alpha+1)}{\Gamma(\alpha)\beta} \int_0^\infty \frac{\beta^{\alpha+1} x^{(\alpha+1)-1} e^{-\beta x}}{\Gamma(\alpha+1)} dx$$

$$= \frac{\Gamma(\alpha+1)}{\Gamma(\alpha)\beta}$$ × 仿佛只与α有关.

$$= \frac{\alpha}{\beta}$$

$$E(x^2) = \frac{1}{\Gamma(\alpha)} \int_0^\infty \beta^\alpha \cdot x^{(\alpha+2)-1} e^{-\beta x} dx$$

$$= \frac{\Gamma(\alpha+2)}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty \frac{\beta^{(\alpha+2)} x^{(\alpha+2)-1} e^{-\beta x}}{\Gamma(\alpha+2)} dx$$

$$= \frac{\alpha(\alpha+1)}{\beta^2}$$

$$Var(x) = E(x^2) - [E(x)]^2$$

$$= \frac{\alpha}{\beta^2}$$

Gamma Distribution. $G(\alpha, \beta)$ or $\Gamma(\alpha, \beta)$, $\alpha, \beta > 0$

$$f_X(x) = \begin{cases} \dfrac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)} & x > 0 \\ 0 & x \leq 0. \end{cases}$$

$$\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} e^{-y} dy.$$

$$\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1) \qquad \Gamma(\alpha) = (\alpha-1)!$$

$$E(x) = \frac{\alpha}{\beta} \qquad Var(x) = \frac{\alpha}{\beta^2}$$

if $x_1, \dots x_n$ are indep. Exponential distribution each with mean β^{-1}, then $\sum_{i=1}^{n} x_i \sim G(n, \beta)$

Beta distribution. $\alpha, \beta > 0$.

$$f_X(x) = \begin{cases} \dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1} & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

Special case. If $\alpha = \beta = 1$ then $f_X(x) = 1$ i.e. $X \sim U(0,1)$

$$E(x) = \frac{\alpha}{\alpha+\beta} \qquad Var(x) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

$$\int_0^1 f_X(x) dx = 1$$

$$\int_0^1 x^{\alpha-1}(1-x)^{\beta-1} = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

$$E(x) = \int \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot x^{(\alpha+1)-1}(1-x)^{\beta-1} dx$$

$$= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot \frac{\Gamma(\alpha+1)\Gamma(\beta)}{\Gamma(\alpha+\beta+1)}$$

$$= \frac{(\alpha+\beta-1)!\,\alpha!}{(\alpha+\beta)!\,(\alpha-\beta)!}$$

$$= \frac{\alpha}{\beta+\beta}$$

$$E(x^2) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot \frac{\Gamma(\alpha+2)\Gamma(\beta)}{\Gamma(\alpha+\beta+2)}$$

$$= \frac{(\alpha+\beta-1)! \cdot (\alpha+1)!}{(\alpha-1)!\,(\alpha+\beta+1)!}$$

$$= \frac{\alpha(\alpha+1)}{(\alpha+\beta+1)(\alpha+\beta)}$$

Summary

$$Var(x) = E(x^2) - [E(x)]^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

DISCRETE BIVARIATE DISTRIBUTIONS

X, Y. two discrete random variables. taking values $\{x_i\}$ and $\{y_j\}$ respectively. Their joint pmf. $P_{X,Y}(x,y)$.

$$\boxed{\begin{aligned} P_{X,Y}(x,y) &= P(\{X=x\}\cap\{Y=y\}) \\ &= P(X=x, Y=y) \end{aligned}}$$

cdf.

$$\boxed{\begin{aligned} F_{X,Y}(x,y) &= P(\{X\le x\}\cap\{Y\le y\}) \\ &= P(X\le x, Y\le y) \end{aligned}}$$

Properties.

① $0 \le P_{X,Y}(x,y) \le 1$

② $\sum_i \sum_j P_{X,Y}(x_i, y_j) = 1.$

③ $$\boxed{\begin{aligned} P_X(x_i) &= P(X=x_i) = \sum_j P_{X,Y}(x_i, y_j) \\ P_Y(y_j) &= P(Y=y_j) = \sum_i P_{X,Y}(x_i, y_j) \end{aligned}}$$ marginal pmf.

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

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

④ If $g(X,Y)$ is a function of (X,Y) then

$$E[g(X,Y)] = \sum_i \sum_j g(x_i, y_j) P_{X,Y}(x_i, y_j)$$

eg: $E[X] = \sum_i \sum_j x_i P_{X,Y}(x_i, y_j).$

Covariance of X Y.

$Cov(X,Y) = E[(X-E(X))(Y-E(Y))]$

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

The correlation coefficient

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}} \qquad -1 \le \rho(X,Y) \le 1$$

$\rho(X,Y)=0 \iff Cov(X,Y)=0.$

Summary

Proof:

$$E[XY] = \sum_i \sum_j x_i y_j \, p_{X,Y}(x_i, y_j)$$
$$= \sum_i \sum_j x_i y_j \, p_X(x_i) \, p_Y(y_j)$$
$$= \left(\sum_i x_i p_X(x_i)\right)\left(\sum_j y_j p_Y(y_j)\right)$$
$$= E[X]\,E[Y]$$

$$\therefore \mathrm{Cov}(X,Y) = E[XY] - E[X]E[Y] = 0$$

Proof:

$$E(X) = \sum_x \sum_y x \, p_{X,Y}(x,y)$$
$$= \sum_y p_Y(y) \cdot \sum x \frac{p_{X,Y}(x,y)}{p_Y(y)}$$
$$= \sum_y p_Y(y) \sum_x x \cdot p_{X|Y}(x|y)$$
$$= \sum_y p_Y(y) \, E[X \mid Y=y].$$

Proof: $p_Z(n) = \sum_{k=0}^{n} p_{X,Y}(k, n-k)$

by the theorem of total probability.

$$p_Z(n) = p(Z=n) = \sum_{k=0}^{n} p(Z=n \mid X=k) \cdot p(X=k)$$
$$= \sum_{k=0}^{n} p(Z=n \text{ and } X=k)$$
$$= \sum_{k=0}^{n} p(X=k, Y=n-k)$$
$$= \sum_{k=0}^{n} p_{X,Y}(k, n-k)$$

如果单纯 $\mathrm{Cov}(X,Y)=0$，无法判断 X,Y 是否 indep.

Independent.

$p_{X,Y}(x,y) = p_X(x)\, p_Y(y)$ for all x and y.

Theorem: If X and Y are independent, then $\mathrm{Cov}(X,Y)=0$

$\mathrm{Cov}(X,Y) \neq 0 \Rightarrow$ X and Y are dependent.

Conditional distribution.

$$p(E \mid F) = \frac{p(E \cap F)}{p(F)}$$

$$p_{X|Y}(x|y) = p(X=x \mid Y=y) = \frac{p(X=x, Y=y)}{p(Y=y)} = \frac{p_{X,Y}(x,y)}{p_Y(y)}$$

Conditional Expectation.

$$E[X \mid Y=y] = \sum_x x \, p_{X|Y}(x|y)$$

Theorem $E(X) = \sum_y p_Y(y) \cdot E[X \mid Y=y]$

因为我求它的时候会先求 $p_{X|Y}(x|y)$.

Summary

Prove. Assum. X, Y discrete.

$E[X+Y] = \sum_i \sum_j (x_i+y_j) P_{X,Y}(x_i, y_j)$

$= \sum_i \sum_j x_i P_{X,Y}(x_i,y_j) + \sum_i \sum_j y_j P_{X,Y}(x_i,y_j)$

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

$E[XY] = \sum_i \sum_j x_i y_j P_{X,Y}(x_i, y_j)$

$= \sum_i \sum_j x_i y_j P_X(x_i) P_Y(y_j)$

$= \sum_i x_i p_X(x_i) \sum_j y_j P_Y(y_j)$

$= E[X] \cdot E[Y]$.

$E[(X+Y)^2] = E(X^2 + 2XY + Y^2)$

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

$Var(X+Y) = E[(X+Y)^2] - E^2[(X+Y)]$

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

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

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

$Z = X + Y$. $n = 0, 1, \cdots$

$$\boxed{P_Z(n) = \sum_{k=0}^{n} P_{X,Y}(k, n-k)}$$

Special case: If X and Y are independent $Z = X + Y$

$$\boxed{P_Z(n) = \sum_{k=0}^{n} P_X(k) P_Y(n-k)}$$

① $E[X+Y] = E[X] + E[Y]$

② X, Y independent

$E[XY] = E[X]E[Y]$

$var(X+Y) = var(X) + var(Y)$

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

$\Downarrow$ generalize

$X_1, X_2, \cdots X_n$ be random variables.

① $E[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} E[X_i]$

② if $X_1, X_2 \cdots X_n$ are independent then

$E[\prod_{i=1}^{n} X_i] = \prod_{i=1}^{n} E[X_i]$ $\quad Var[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} var[X_i]$

In particular. if X_1, X_2, X_n are indept. and identically distributed with $E[X] = \mu$ and $Var(X) = \sigma^2$. then.

$E[\sum_{i=1}^{n} X_i] = n\mu$ and $Var(\sum_{i=1}^{n} X_i) = n\sigma^2$

Summary

Probability Generating Functions.

pgf:

$$P_X(s) = E[s^X] = \sum_{k=0}^{\infty} P_X(k) s^k.$$

Note: $P_X(s)$ is defined for all $s \in [-1, 1]$

$$\sum_{k=0}^{\infty} |P_X(k) s^k| \le \sum_{k=0}^{\infty} P_X(k) = 1$$

Let $P_X^{(n)}(s)$ denote the n^{th} derivative of $P_X(s)$

$$\boxed{P_X^{(n)}(1) = E[X(X-1)\cdots(X-n+1)]} \quad (n = 1, 2, \cdots)$$

Proof:

$$P_X^{(n)}(s) = \frac{d^n}{ds^n}\left(\sum_{k=0}^{\infty} P_X(k) s^k\right)$$

$$= \sum_{k=n}^{\infty} P_X(k) k(k-1)\cdots(k-n+1) s^{k-n}$$

$$\therefore P_X^{(n)}(1) = \sum_{k=n}^{\infty} P_X(k) k(k-1)\cdots(k-n+1)$$

$$= E[X(X-1)\cdots(X-n+1)]$$

Special case.

① $E(X) = P_X'(1)$

② $Var(X) = P_X''(1) + P_X'(1) - [P_X'(1)]^2$

求一次导. 少1

求n次导. 少n.

Proof:

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

$$= E[X(X-1)] + E(X) - [E(X)]^2$$

$$= P_X''(1) + P_X'(1) - [P_X'(1)]^2$$

$$P_X^{(n)}(s) = \sum_{k=n}^{\infty} P_X(k) k(k-1)\cdots(k-n+1) s^{k-n}$$

$$\Rightarrow P_X^{(n)}(0) = P_X(n)\, n!$$

$Z = X + Y \quad P_Z(s) = P_X(s) P_Y(s)$

Proof:

$$P_Z(s) = E[s^Z] = E[s^{X+Y}]$$

$$= E[s^X \cdot s^Y]$$

$$= E[s^X] E[s^Y]$$

$$= P_X(s) \cdot P_Y(s)$$

$Z = X_1 + X_2 + \cdots + X_n \quad P_Z(s) = \prod_{i=1}^{n} P_{X_i}(s)$

Summary

In other words, the CDF of $\frac{S_n - n\mu}{\sigma\sqrt{n}}$ tends to that of a standard normal random variables as $n \to \infty$. This implies $S_n \approx N(n\mu, n\sigma^2)$

Central Limit Theorem. Let $X_1, X_2 \cdots X_n$ be independent and identically distributed (iid) random variables, each with mean μ and variance σ^2. Let $S_n = X_1 + X_2 + \cdots + X_n$.

Then for all $x \in \mathbb{R}$

$$P\left[\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq x\right] \to P[Z \leq x] \text{ as } n \to \infty$$

where $Z \sim N(0,1)$

The Sample mean $\bar{X}$

$$\bar{X} = \frac{S_n}{n} = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

$$E[\bar{X}] = E\left[\frac{S_n}{n}\right] = \frac{1}{n} E[\overset{n\mu}{X_1 + X_2 + \cdots + X_n}] = \mu$$

$$\text{var}[\bar{X}] = \text{var}\left[\frac{S_n}{n}\right] = \frac{1}{n^2} \overset{n\sigma^2}{\text{Var}(S_n)} = \frac{\sigma^2}{n}$$

So. $S_n \approx N(n\mu, n\sigma^2)$

$\bar{X} \approx N(\mu, \sigma^2/n)$

Normal Approximation to Binomial Distribution

$X = X_1 + X_2 + \cdots + X_n$ $\quad E[X_i] = p \quad \text{var}(X_i) = p(1-p)$

Restrictions: $\boxed{n \geq 20 \quad np \geq 5 \quad n(1-p) \geq 5}$

Then $X \approx N(np, np(1-p))$

Continuity correction:

$P(X = x) \approx P(x - 0.5 < Y < x + 0.5)$

where $Y \sim N(np, np(1-p))$

Summary

$$\lim_{n \to \infty} F_n(x) = \lim_{n \to \infty} \left\{ \frac{\sum_i^n X_i - n\mu}{\sqrt{n}\,\sigma} \leq x \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{x^2}{2}} dx$$