$E[X^k] = M^{(k)}(0)$

Table of Discrete and Continuous distributions

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

$M(t) = E[e^{tX}]$, $M(0) = 1$

Distribution	Type	Mass/density function $f(x)$ (PMF / PDF)	Mean μ ($E[x]$)	Variance σ^2	Moment Generating Function
UNIFORM(n)	D	$1/n$, for $x = 1, 2, \ldots, n$	$(n+1)/2$	$(n^2-1)/12$	$\frac{e^{at} - e^{(b+1)t}}{(b-a+1)(1-e^t)}$, $t \in \mathbb{R}$
UNIFORM(a,b)	C	$\frac{1}{b-a}$, for $x \in [a,b]$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{e^{tb} - e^{ta}}{tb - ta}$, $t \in \mathbb{R}$
BERNOULLI(p)	D	$f(0) = 1-p, f(1) = p$	p	$p(1-p)$	$q + pe^t$, $t \in \mathbb{R}$
BINOMIAL(n,p)	D	$\binom{n}{x} p^x (1-p)^{n-x}$, for $x = 0, 1, \ldots, n$	np	npq	$(q + pe^t)^n$, $t \in \mathbb{R}$
GEOMETRIC(p)	D	$q^{x-1}p$, for $x = 1, 2, \ldots$	$1/p$	$(1-p)/p^2$	$\frac{p}{1-qe^t}$, $t < -\log q$
POISSON(λt)	D	$\frac{1}{x!}(\lambda t)^x e^{-\lambda t}$, for $x = 0, 1, \ldots$	λt	λt	$e^{\lambda(e^t - 1)}$, $t \in \mathbb{R}$
HYPERGEOMETRIC(N,M,n)	D	$\frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$, for $x = 0, 1, \ldots, n$	np	$np(1-p)$	Too Complex
EXPONENTIAL(λ)	C	$\lambda e^{-\lambda x}$, for $x \in [0, \infty)$	$1/\lambda$	$1/\lambda^2$	$\frac{\lambda}{\lambda - t}$, $t < \lambda$
GAMMA(α, β)	C	$\frac{x^{\alpha-1}e^{-x/\beta}}{\beta^\alpha \Gamma(\alpha)}$, for $x \in [0, \infty)$	$\alpha\beta$	$a\beta^2$	$\left(\frac{1}{1-\beta t}\right)^\alpha$, $t < \frac{1}{\beta}$
NORMAL(μ, σ^2) $N(0,1)$: $E(X^n) = \begin{cases} 0 & ; n \text{ even} \\ (n-1)(n-3)\cdots 3 \cdot 1 & ; n \text{ odd} \end{cases}$	C	$\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$, for $x \in \mathrm{R}$	μ	σ^2	$e^{\mu t + \frac{\sigma^2 t^2}{2}}$, $t \in \mathbb{R}$
Degenerate (a)	D	$P(X = x) = \begin{cases} 1 & ; x = a \\ 0 & ; x \neq a \end{cases}$	a	0	e^{at}, $t \in \mathbb{R}$
Multinomial $(m, p_1, \ldots, p_{k-1})$		$P(X_1 = x_1, \ldots, X_{k-1} = p_{k-1}) = \binom{m}{x_1, \ldots, x_k} \prod_{j=1}^{k} p_j^{x_j}$, $\binom{m}{x_1, \ldots, x_k} = \frac{m!}{x_1! \cdots x_k!}$			

$$\begin{cases} E\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i E[X_i] \\ Var\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i^2 Var(X_i) + \sum_{i,j}^{n} a_i a_j \overbrace{Cov(X_i, X_j)}^{=0 \text{ if } X_i \perp\!\!\!\perp X_j} \end{cases}$$

$\underline{X_i \text{ IID}}$: $\begin{cases} E\left(\sum_{i=1}^{n} X_i\right) = n\mu \\ Var\left(\sum_{i=1}^{n} X_i\right) = n\sigma^2 \end{cases}$ $\begin{cases} E\left(\sum_{i=1}^{n} \frac{X_i}{n}\right) = \mu \\ Var\left(\sum_{i=1}^{n} \frac{X_i}{n}\right) = \frac{\sigma^2}{n} \end{cases}$

$E(X_i) = \mu$
$Var(X_i) = \sigma^2$

$\underline{X_i \text{ IID}}$: $M_{\sum a_i X_i}(t) = \prod_{i=1}^{n} M_{X_i}(t a_i)$

$$\begin{cases} X_n \xrightarrow{P} X \iff \forall \epsilon > 0 : P(|X_n - X| > \epsilon) = 0 \text{ as } n \to \infty \\ X_n \xrightarrow{d} X \iff F_{X_n}(x) \to F_X(x) \ \forall x \in \mathbb{R} \text{ as } n \to \infty \end{cases}$$

↳ a continuity pt of F_X

$X_n \xrightarrow{d} c \Rightarrow X_n \xrightarrow{P} c$
$X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{d} X$

$\underline{\text{WLOLN}}$: (X_n) IID, $\begin{cases} E(X_i) = \mu < \infty \\ Var(X_i) = \sigma^2 < \infty \end{cases}$, $\underline{E(X_i^2) < \infty}$:

$$\overline{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} \mu = E[X_i] \text{ as } n \to \infty$$

$\underline{\text{Slutsky}}$: $\begin{cases} X_n \xrightarrow{d} X \\ Y_n \xrightarrow{P} c \in \mathbb{R} \end{cases} \Rightarrow \begin{cases} X_n \pm Y_n \xrightarrow{d} X \pm c \\ X_n Y_n \xrightarrow{d} cX \\ \underline{c \neq 0}: \frac{X_n}{Y_n} \xrightarrow{d} \frac{X}{c} \end{cases}$

$\underline{\text{Note}}$: $c = 0 \Rightarrow X_n Y_n \xrightarrow{P,d} 0$

$\underline{\text{CLT}}$: $\begin{cases} Z_n = \frac{\overline{X_n} - \mu}{\sigma/\sqrt{n}} \xrightarrow{d} Z \sim N(0,1) \\ Z_n = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n\sigma^2}} \xrightarrow{d} Z \sim N(0,1) \end{cases}$

$\underline{\text{IID}}$: $\begin{cases} E(X_i) = \mu < \infty \\ Var(X_i) = \sigma^2 < \infty \end{cases}$

$\Rightarrow \begin{cases} \overline{X_n} \approx N\left(\mu, \frac{\sigma^2}{n}\right) \\ \sum_{i=1}^{n} X_i \approx N(n\mu, n\sigma^2) \end{cases}$ for n large.

- $E[g(X)] = \begin{cases} \sum_{x \in S_X} g(x) \cdot P(X=x) \\ \int_{-\infty}^{\infty} g(x) f_X(x)\,dx \end{cases}$

$E(|X|^t) < \infty \Rightarrow \begin{cases} E(|X|^s) < \infty \quad \forall s \le t \\ \lim_{n \to \infty} n^t P(|X| > n) = 0 \end{cases}$

$\begin{cases} E[aX+b] = a\,E[X] + b \\ Var[aX+b] = a^2 Var[X] \end{cases}$

- p^{th} Quantile of X: $\begin{cases} P(X \le x_p) \ge p \\ P(X \ge x_p) \le 1-p \end{cases}$, i.e. $p \le F(x_p) \le p + F(X = x_p)$

- $E(h(X)) < \infty \Rightarrow P(h(X) \ge M) \le \frac{E[h(X)]}{M}, \forall M > 0$ — $h(X) \ge 0$!!!

- $E(|X|^k) < \infty \Rightarrow P(|X| \ge M) \le \frac{E[|X|^k]}{M^k}, \forall M > 0$ (Markov)

- $\left.\begin{matrix} E(X) = \mu, \; Var(X) = \sigma^2 \\ E(X^2) < \infty \end{matrix}\right\} \Rightarrow P(|X - \mu| \ge k\sigma) \le \frac{1}{k^2}, \forall k > 0$ (Chebychev-Biênayme)

- $E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i]$ if $E[|X_i|] < \infty$

- $F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(x,y)\,dy\,dx \iff f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y}$

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

- $F_X(x) = \lim_{y \to \infty} F(x,y)$; $F_Y(y) = \lim_{x \to \infty} F(x,y)$

$\begin{cases} f_X(x) = \frac{\partial F_X(x)}{\partial x} \\ f_Y(y) = \frac{\partial F_Y(y)}{\partial y} \end{cases}$

- Cond PMF = $\frac{\text{Joint PMF}}{\text{Marginal PMF} > 0}$: $P(X = x_i \mid Y = y_j) = \frac{p_{ij}}{p_{\cdot j}} = \frac{P(X = x_i, Y = y_j)}{P(Y = y_j)}$

- Cond PDF = $\frac{\text{Joint PDF}}{\text{Marginal PDF} > 0}$: $f_{X|Y=y}(x) = \frac{f(x,y)}{f_Y(y)}$

- Cond CDF : $F_{X|Y=y}(x) = \int_{-\infty}^{x} f_{X|Y=y}(t)\,dt$; $F_{X|Y=y_j}(x_i) = \sum_{k=1}^{i} \frac{p_{kj}}{p_{\cdot j}}$

- $\begin{cases} E[X] = E[E[X|Y]] \\ E[g(X)] = E[E[g(X)|Y]] \end{cases}$

- $\begin{cases} Var[X|Y=y] = E[(X - \mu_{X|Y=y})^2 \mid Y=y] = E[X^2|Y=y] - (E[X|Y=y])^2 \\ Var[X] = Var[E(X|Y)] + E[Var(X|Y)] \end{cases}$

$E[X|Y=y] = \mu_{X|Y=y}$

Indep = $\mathbb{X} \perp\!\!\!\perp Y \iff F_{\mathbb{X},Y}(x,y) = F_{\mathbb{X}}(x)\, F_Y(y) \iff \begin{cases} P(\mathbb{X}=x, Y=y) = P(\mathbb{X}=x)\,P(Y=y) \\ f_{\mathbb{X},Y}(x,y) = f_{\mathbb{X}}(x) f_Y(y) \end{cases}$

$$X \perp\!\!\!\perp Y \implies \begin{cases} E(\mathbb{X}Y) = E(\mathbb{X})E(Y) \\ E(g_1(\mathbb{X})g_2(Y)) = E(g_1(\mathbb{X}))E(g_2(Y)) \end{cases}$$

Identically dist = $F_{\mathbb{X}_i}(x) = F_{\mathbb{X}_j}(x) \quad \forall i,j \in \{1,\dots,n\},\ \forall x \in \mathbb{R} \rightsquigarrow F_{\mathbb{X}}(x) = F_Y(x)\ \forall x \in \mathbb{R}$.

$\hookrightarrow \mathbb{X}_i \stackrel{d}{=} \mathbb{X}_j$

Transfo = $\vec{Y} = g(\vec{\mathbb{X}})$
$$\begin{cases} f_{Y_1,\dots,Y_n}(y_1,\dots,y_n) = f_{\mathbb{X}_1,\dots,\mathbb{X}_n}\left(g_1^{-1}(y_1,\dots,y_n),\dots,g_n^{-1}(y_1,\dots,y_n)\right) \cdot |J| \\ J = \det\begin{pmatrix} \frac{\partial \mathbb{X}_1}{\partial y_1} & \cdots & \frac{\partial \mathbb{X}_1}{\partial y_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial \mathbb{X}_n}{\partial y_1} & \cdots & \frac{\partial \mathbb{X}_n}{\partial y_n} \end{pmatrix} \neq 0 \end{cases}$$

MGF = $M_{\vec{\mathbb{X}}}(t_1,\dots,t_n) = E\left[e^{t_1\mathbb{X}_1+\dots+t_n\mathbb{X}_n}\right] = \begin{cases} \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} e^{t_1 x_1+\dots+t_n x_n} f(x_1,\dots,x_n)\,dx_1\cdots dx_n \\ \sum_{\vec{x}\in S_{\vec{X}}} e^{t_1 x_1+\dots+t_n x_n} P(\mathbb{X}_1=x_1,\dots,\mathbb{X}_n=x_n) \end{cases}$

$\hookrightarrow M_{\vec{\mathbb{X}}}(0,0,\dots,0) = 1 \quad , \quad M_{\vec{\mathbb{X}}}(0,\dots,t_i,\dots,0) = M_{\mathbb{X}_i}(t_i)$

- $E[\mathbb{X}_1^m \mathbb{X}_2^n] = \dfrac{\partial^{m+n} M_{\vec{\mathbb{X}}}(t_1,t_2)}{\partial t_1^m\, \partial t_2^n}$

- $\vec{\mathbb{X}}_i \perp\!\!\!\perp \implies M_{\vec{\mathbb{X}}}(t_1,\dots,t_n) = \prod_{i=1}^n M_{\mathbb{X}_i}(t_i)$
- $\mathbb{X}_i$ IID $\implies M_{\vec{\mathbb{X}}}(t_1,\dots,t_n) = \prod_{i=1}^n M_{\mathbb{X}_1}(t_i) \rightsquigarrow M_{\vec{\mathbb{X}}}(t_1,\dots,t_n) = \left[M_{\mathbb{X}_1}(t_1)\right]^n$

Cauchy Schwartz : $0 \le E(|\mathbb{X}Y|) \le \sqrt{E(\mathbb{X}^2)E(Y^2)}$
$0 \le E(|\mathbb{X}Y|) \le (E(|\mathbb{X}|^p))^{1/p}(E(|Y|^q))^{1/q} \quad , \frac{1}{p}+\frac{1}{q}=1$

$\to Cov(\mathbb{X},Y) = \sigma_{\mathbb{X},Y} = E[(\mathbb{X}-\mu_X)(Y-\mu_Y)] = E[\mathbb{X}Y] - E[\mathbb{X}]E[Y]$ $\begin{cases} >0 : \text{Concordant} \\ <0 : \text{discordant} \\ =0 : \text{uncorr} \end{cases}$

$\hookrightarrow$ indep $\implies \sigma_{\mathbb{X}Y} = 0$

- $Cov(a\mathbb{X}+b, cY+d) = ac\,Cov(\mathbb{X},Y)$
- $\sigma_{\mathbb{X}\mathbb{X}} = \sigma_X^2$

$\to Corr(\mathbb{X},Y) = \rho_{\mathbb{X},Y} = \dfrac{\sigma_{\mathbb{X},Y}}{\sigma_{\mathbb{X}}\cdot\sigma_Y} = \dfrac{Cov(\mathbb{X},Y)}{\sqrt{Var(\mathbb{X})\,Var(Y)}}$ $\rightsquigarrow$ measure Linear relat° between $\mathbb{X},Y$:
$Y = a\mathbb{X}+b \iff \rho_{\mathbb{X},Y} = \begin{cases} +1 ; a>0 \\ -1 ; a<0 \end{cases}$

- $|\rho_{\mathbb{X},Y}| \le 1 \quad , \rho_{\mathbb{X}\mathbb{X}} = 1$
- $\rho_{a\mathbb{X}+b,\, cY+d} = \mathrm{Sgn}(ac)\,\rho_{\mathbb{X},Y}$