Lec 6

6.1 Joint distribution functions

$F_X(x)=F(x,\infty)$, $F_Y(y)=F(\infty,y)$

$P\{X>x, Y>y\}=1-F_X(x)-F_Y(y)+F(x,y)$

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

$P_X(x)=P\{X=x\}=\sum_{y:p(x,y)>0} p(x,y)$ (离散)

prop: $p(x,y)\ge 0$, $\sum_x\sum_y p(x,y)=1$ (连续同理)

连续. $P\{(X,Y)\in D\}=\iint_{(x,y)\in D} f(x,y)\,dxdy$

$F(x,y)=P\{X\le x, Y\le y\}=\int_{-\infty}^{y}\int_{-\infty}^{x} f(u,v)\,dudv$

$f(x,y)=\frac{\partial^2}{\partial x\partial y}F(x,y)$

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

给定 marginal distribution, 不能确定 joint distribution

6.2 Independent random variables

Independent: $P\{X=x, Y=y\}=P\{X=x\}P\{Y=y\}$

or $p(x,y)=p_X(x)p_Y(y)$ $F(a,b)=F_X(a)F_Y(b)$

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

prop.

X, Y independent $\Longleftrightarrow f(x,y)=h(x)g(x)$

$-\infty<x,y<\infty$

6.3 Sum of independent random variable

X, Y independent.

$F_{X+Y}(a)=\int_{-\infty}^{\infty}F_X(a-y)f_Y(y)\,dy$

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

$X_1, X_2, \cdots, X_n\sim U(0,1)$ i.i.d, $F_n\sim X_1+X_2+\cdots+X_n$

Then $F_n(x)=\frac{x^n}{n!}$ $(0\le x\le 1)$

Let $N=\min\{n: X_1+\cdots+X_n>1\}$

Then $E[N]=e$

~~Gamma~~ Gamma distribution

$X\sim Gamma(k,\lambda)$

$f(x)=\frac{\lambda^k x^{k-1}e^{-\lambda x}}{\Gamma(k)}$, $0<x<\infty$

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

Remark

$Exp(\lambda)=Gamma(1,\lambda)$. if $X_1,\cdots,X_n\sim Exp(\lambda)$, then $X_1+\cdots+X_n\sim Gamma(n,\lambda)$

X_1	$Ber(p)$	$Binom(n,p)$	$Poisson(\lambda_1)$
X_2	$Ber(p)$	$Binom(m,p)$	$Poisson(\lambda_2)$
X_1+X_2	$Binom(2,p)$	$Binom(n+m,p)$	$Poisson(\lambda_1+\lambda_2)$
	$Exp(\lambda)$	$\Gamma(k_1,\lambda)$	$N(\mu_1,\sigma_1^2)$
	$Exp(\lambda)$	$\Gamma(k_2,\lambda)$	$N(\mu_2,\sigma_2^2)$
	$\Gamma(2,\lambda)$	$\Gamma(k_1+k_2,\lambda)$	$N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2)$

χ^2-distribution

$Z_1,\cdots,Z_n\sim N(0,1)$ ~~indep~~ i.i.d, $Y=\sum_{i=1}^n Z_i^2$

have the χ^2-distribution with n freedom

Normal distribution

$X_1,\cdots,X_n\sim N(\mu_i,\sigma_i^2)$ independent

$\sum_{i=1}^n a_iX_i\sim N\left(\sum_{i=1}^n a_i\mu_i, \sum_{i=1}^n a_i^2\sigma_i^2\right)$

6.4 Conditional distribution.

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

$F_{X|Y}(x|y)=P\{X\le x|Y=y\}=\sum_{u\le x}p_{X|Y}(u|y)$

$f_{X|Y}(x,y)=\frac{f(x,y)}{f_Y(y)}$ for all y such that $f_Y(y)>0$

Joint probability distribution of $g(X,Y)$

$U=g(X,Y)$, $V=h(X,Y)$

$$J(x,y)=\begin{vmatrix}\frac{\partial g}{\partial x} & \frac{\partial g}{\partial y}\\ \frac{\partial h}{\partial x} & \frac{\partial h}{\partial y}\end{vmatrix}\ne 0$$

$f_{U,V}(u,v)=f_{X,Y}(x,y)\,|J(x,y)|^{-1}$ (x,y 用u,v表示)

Lec 7 Expectation

7.1 Expectation of sums of random variables

$E[g(X,Y)]=\sum_y\sum_x g(x,y)p(x,y)$

$E[g(X,Y)]=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f(x,y)\,dxdy$

$E[aX+bY]=aE[X]+bE[Y]$

If $X\ge Y$ a.s. then $E[X]\ge E[Y]$

Sample mean: $\bar{X}=\sum_{i=1}^n\frac{X_i}{n}$ $E[\bar{X}]=\mu$ ~~$E[\bar{X}]=\mu$~~

Boole's inequality: $P\left(\bigcup_{i=1}^n A_i\right)\le\sum_{i=1}^n P(A_i)$

7.2 Covariance

X, Y independent, $E[g(X)h(Y)]=E[g(X)]E[h(Y)]$

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

if X, Y independent, then $Cov(X,Y)=0$

Prop. $Cov(X,Y)=Cov(Y,X)$ 推出不相关.

$Cov(X,X)=Var(X)$

$Cov(aX,bY)=ab\,Cov(X,Y)$

$Cov(X_1+X_2,Y)=Cov(X_1,Y)+Cov(X_2,Y)$

$Cov(a_1X_1+\cdots+a_nX_n, b_1Y_1+\cdots+b_mY_m)=\sum_{i=1}^n\sum_{j=1}^m a_ib_j\,Cov(X_i,Y_j)$

Sample variance: $S^2=\frac{1}{n-1}\sum_{i=1}^n(X_i-\bar{X})^2$

$Var(\bar{X})=Cov(\bar{X},\bar{X})=\frac{\sigma^2}{n}$ $E[S^2]=\sigma^2$

$Var\left(\sum_{i=1}^n X_i\right)=\sum_{i=1}^n Var(X_i)+2\sum\sum_{i<j}Cov(X_i,Y_j)$

Correlation.

$Cor(X,Y)=\rho(X,Y)=\frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$

衡量线性关系 $\rho\in[-1,1]$

$|E[XY]|\le\sqrt{E[X^2]E[Y^2]}$, 当$X=cY$时取等.

$Cov(X,Y)=1\Longleftrightarrow P\left\{\frac{X-E[X]}{SD(X)}=\frac{Y-E[Y]}{SD(Y)}\right\}=1$

$Cov(X,Y)=-1\Longleftrightarrow P\left\{\frac{X-E[X]}{SD(X)}=-\frac{Y-E[Y]}{SD(Y)}\right\}=1$

7.3 Conditional expectation

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

$$E[X|Y=y]=\begin{cases}\int_{-\infty}^{\infty}x f_{X|Y}(x,y)\,dx, & f_Y(y)>0\\ 0\end{cases}$$

if $E[X|Y=y]=h(y)$, then we write $E[X|Y]=h(Y)$

- $E[X]=E[E[X|Y]]$

$E\left[\sum_{i=1}^n X_i\middle|Y=y\right]=\sum_{i=1}^n E[X_i|Y=y]$

$$E[g(X)|Y=y]=\begin{cases}\sum_x g(x)p_{X|Y}(x|y)\\ \int_{-\infty}^{\infty}g(x)p_{X|Y}(x|y)\,dx\end{cases}$$

最小二乘法:

$h=\arg\min E[\{Y-g(X)\}^2]$

$h(x)=E[Y|X=x]$ (regression function)

Total expectation.

$E[X]=E[E[X|Y]]=\int_{-\infty}^{\infty}E[X|Y=y]f_Y(y)\,dy$

$$P(E)=\begin{cases}\sum_y P(E|Y=y)P_Y(y)\\ \int_{-\infty}^{\infty}P(E|Y=y)f_Y(y)\,dy\end{cases}$$

Conditional Variance.

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

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

7.4 Moment generating function.

$M(t)=E[e^{tX}]$ for all $t\in\mathbb{R}$.

$M'(t)=E[Xe^{tX}]$, $\therefore M'(0)=E[X]$, $M''(0)=E[X^2]$

$X\sim Binom(n,p)$, then $M(t)=(pe^t+1-p)^n$

$X\sim Exp(\lambda)$, then $M(t)=\frac{\lambda}{\lambda-t}$ for $t<\lambda$

$X\sim N(\mu,\sigma^2)$, then $M(t)=e^{\mu t+\frac{\sigma^2t^2}{2}}$

X, Y independent, $M_{X+Y}(t)=M_X(t)M_Y(t)$

mgf uniquely determine the distribution.

Joint: $M(s,t)=E[e^{sX+tY}]$

$M_X(s)=M(s,0)$

independent $\Longleftrightarrow M(s,t)=M_X(s)M_Y(t)$

ex: $X, Y\sim N(\mu,\sigma^2)$ i.i.d. find joint $X+Y$, $X-Y$

$E[e^{s(X+Y)+t(X-Y)}]=M(s+t)M(s-t)$
$=e^{2\mu s+\sigma^2s^2}e^{\sigma^2t^2}$

$X+Y\sim N(2\mu,2\sigma^2)$, $X-Y\sim N(0,2\sigma^2)$ independent

Lec 8 Multivariate normal distribution.

8.1 Bivariate

Standard bivariate normal distribution

$f_X(x)=\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$ $f_Y(y)=\frac{1}{\sqrt{2\pi}}e^{-y^2/2}$ $X\sim N(0,1)$ $Y\sim N(0,1)$

$f(x,y)=\frac{1}{2\pi}e^{-\frac{x^2}{2}-\frac{y^2}{2}}$

$f_{X,Y}(x,y)=\frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}$ Parameter: $(\mu_x,\mu_y,\sigma_x^2,\sigma_y^2,\rho)$

$\exp\left(-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_x)^2}{\sigma_x^2}-2\rho\frac{(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y}+\frac{(y-\mu_y)^2}{\sigma_y^2}\right]\right)$

Z-score:

$X^*=\frac{X-\mu}{\sigma}$, $\mu=E[X]$, $\sigma^2=Var(X)$

if $X\sim N(\mu,\sigma^2)$, then $X^*\sim N(0,1)$

$\rho=Cor(X,Y)=Cor(X^*,Y^*)=Cov(X^*,Y^*)$

if $U\sim N(0,1)$, $V\sim N(0,1)$

let $X=U$, $Y=\rho U+\sqrt{1-\rho^2}V$ (linear transformation)

(X,Y) follows $(0,0,1,1,\rho)$

If X, Y follows b.n.d $(\mu_x,\mu_y,\sigma_x^2,\sigma_y^2,\rho)$

then $X\sim N(\mu_x,\sigma_x^2)$, $Y\sim N(\mu_y,\sigma_y^2)$

$Cor(X,Y)=\rho$ $Cov(X,Y)=\rho\sigma_x\sigma_y$

Conditional distribution

$Y|X=x \sim N(\mu_Y+\frac{\rho\sigma_Y}{\sigma_X}(x-\mu_X),(1-\rho^2)\sigma_Y^2)$

or, $Y^*|X^*=x^* \sim N(\rho x^*, 1-\rho^2)$

$E[Y|X=x]=\mu_Y+\frac{\rho\sigma_Y}{\sigma_X}(x-\mu_X)$,

$Var(Y|X=x)=(1-\rho^2)\sigma_Y^2$

Linear regression $h(x)=\mu_Y+\rho\frac{\sigma_Y}{\sigma_X}(x-\mu_X)$

If X and Y bivariate normal and uncorrelated, then they are independent.

X,Y follow $(\mu_X,\mu_Y,\sigma_X^2,\sigma_Y^2,\rho) \Longleftrightarrow$

$aX+bY\sim N(a\mu_X+b\mu_Y, a^2\sigma_X^2+2ab\rho\sigma_X\sigma_Y+b^2\sigma_Y^2)$

([illegible] b.n.d)

prop of b.n.d:

1. the marginal distribution normal
2. the conditional distribution normal
3. the conditional expectation is linear of y or x

8.2 Multivariate normal distribution

$X=(X_1,\cdots,X_p)$ with mean vector μ and cov matrix Σ

pdf: $f_X(x)=\frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}}\exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right)$

denote $X\sim N(\mu,\Sigma)$

$\mu=\begin{pmatrix}\mu_1\\ \vdots\\ \mu_p\end{pmatrix}\in\mathbb{R}^p$, $\Sigma=\begin{pmatrix}\sigma_{11}&\cdots&\sigma_{1p}\\ \vdots&&\vdots\\ \sigma_{p1}&\cdots&\sigma_{pp}\end{pmatrix}\in\mathbb{R}^{p\times p}$

standard MND:

$\mu=\vec{0}$, $\Sigma=I_p$, $X\sim N(\vec{0},I_p)$

prop: 1. linear combination of X is normal

2. marginal ... (同上)

$\int_{\mathbb{R}^p} f_X(x)dx=1$

MGF of $N(\mu,\Sigma)$:

$M(t)=\exp\{\mu^Tt+\frac{1}{2}t^T\Sigma t\}$, $t\in\mathbb{R}^p$.

Subvector of X, $\tilde{X}=(X_{k_1},\cdots,X_{k_r})^T$

follows $N(\tilde{\mu},\tilde{\Sigma})$, $\tilde{\mu}=\begin{pmatrix}\mu_{k_1}\\ \vdots\\ \mu_{k_r}\end{pmatrix}$, $\tilde{\Sigma}=\begin{pmatrix}\sigma_{k_1,k_1}&\cdots&\sigma_{k_1,k_r}\\ \vdots&&\vdots\\ \sigma_{k_r,k_1}&\cdots&\sigma_{k_r,k_r}\end{pmatrix}$

The marginal distribution of X_j is $N(\mu_j,\sigma_{jj})$

... of (X_j,X_k) is $N\left(\begin{pmatrix}\mu_j\\ \mu_k\end{pmatrix},\begin{pmatrix}\sigma_{jj}&\sigma_{jk}\\ \sigma_{kj}&\sigma_{kk}\end{pmatrix}\right)$

$\mu_j=E[X_j]$, $\sigma_{jj}=Var(X_j)$, $\sigma_{jk}=Cov(X_j,X_k)$

$X_1,\cdots,X_p$ independent $\Leftrightarrow \sigma_{jk}=0$

X_1,X_2 ... $\Leftrightarrow \Sigma_{12}=0$

Linear transformation.

for $a\in\mathbb{R}^p$

$X\sim N(\mu,\Sigma)\Longleftrightarrow a^TX\sim N(\sum_{j=1}^p a_j\mu_j,\sum_{j=1}^p\sum_{k=1}^p a_ja_k\sigma_{jk})$

for any $C\in\mathbb{R}^{r\times p}$, $CX\sim N(C\mu, C\Sigma C^T)$

$\exists$ orthogonal transformation U, such that

$UX\sim N(U\mu,\Lambda)$, $\Lambda=\begin{pmatrix}\lambda_1&&\\&\lambda_2&\\&&\ddots\ \lambda_p\end{pmatrix}$

UX components are independent.

求特征值. $\det(A-\lambda I)=0$

$(X-\mu)^T\Sigma^{-1}(X-\mu)\sim\chi_p^2$

if $X=\begin{pmatrix}X_1\\ X_2\end{pmatrix}$ follow a p-variate $N\left(\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix},\begin{pmatrix}\Sigma_{11}&\Sigma_{12}\\ \Sigma_{21}&\Sigma_{22}\end{pmatrix}\right)$

then $X_2|X_1\sim N(\mu_2+\Sigma_{21}\Sigma_{11}^{-1}(X_1-\mu_1),\Sigma_{22}-\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12})$

Fisher's Lemma:

Let $X_1,\cdots,X_n$ be i.i.d $N(\mu,\sigma^2)$

$\bar{X}=\frac{1}{n}\sum_{i=1}^n X_i$, $\hat{\sigma}_n^2=\frac{1}{n-1}\sum_{i=1}^n(X_i-\bar{X})^2$

Then (i) $\bar{X}$ and $\hat{\sigma}_n^2$ are independent

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

(iii) $(n-1)\hat{\sigma}_n^2/\sigma^2\sim\chi_{n-1}^2$

Lec 9 Limit theorems

Markov's inequality

$X\geq0$, $P\{X\geq a\}\leq\frac{E[X]}{a}$ for all $a>0$

Chebyshev's inequality

X has finite μ and σ^2, then for any $a>0$

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

If $Var(X)=0$, then $P\{X=E[X]\}=1$

$\boxed{X_n\xrightarrow{P}X}$ as $n\to\infty$: for every $\varepsilon>0$, $P\{|X_n-X|>\varepsilon\}\to0$

$X_n\pm Y_n\xrightarrow{P}a\pm b$, $X_nY_n\xrightarrow{P}ab$, if $b\neq0$, $\frac{X_n}{Y_n}\xrightarrow{P}\frac{a}{b}$

r阶收敛: $E[|X|^r], E[|X_n|^r]<\infty$, $E[|X_n-X|^r]\to0$, then $X_n\xrightarrow{P}X$

(辛钦) Weak law of large numbers

$\xi_1,\xi_2,\cdots$ i.i.d, $\mu=E[\xi_1]$, then $\bar{\xi}_n\xrightarrow{P}\mu$ as $n\to\infty$ (依概率收敛), in other words, for any $\varepsilon>0$, $P\{|\frac{1}{n}\sum_{i=1}^n\xi_i-\mu|>\varepsilon\}\to0$ as $n\to\infty$

$x\in\bigcap_{\varepsilon>0}\bigcup_{N\geq1}\bigcap_{n\geq N}\{x:|f_n(x)-f(x)|\leq\varepsilon\}$ (逐点收敛)

$\boxed{X_n\xrightarrow{a.s.}X}$ as $n\to\infty$: $P\{\lim_{n\to\infty}X_n=X\}=1$ (almost sure)

i.o. infinitely often f.o. finitely often (A_n sequence)

$\{A_n\ i.o.\}=\bigcap_{m=1}^\infty\bigcup_{n=m}^\infty A_n=\lim_{m\to\infty}\underbrace{\bigcup_{n=m}^\infty A_n}_{\sup A_n}=\limsup_{n\to\infty}A_n$

$\{A_n\ f.o.\}=\bigcup_{m=1}^\infty\bigcap_{n=m}^\infty A_n^c=\lim_{m\to\infty}\underbrace{\bigcap_{n=m}^\infty A_n^c}_{\inf A_n^c}=\liminf_{n\to\infty}A_n$

prop: for any $\varepsilon>0$

$X_n\xrightarrow{a.s.}X\Longleftrightarrow \cancel{P\{|X_n-X|>\varepsilon\}}\ P\{|X_n-X|>\varepsilon, i.o.\}=0$

If $\sum_{n=1}^\infty P(A_n)<\infty$, then $P\{A_n, i.o.\}=1$

If $X_n\xrightarrow{a.s.}X$, then $X_n\xrightarrow{P}X$

If $X_n\xrightarrow{P}X$, $\exists\{n_k\}$ s.t. $X_{n_k}\xrightarrow{a.s.}X$

Strong law of large numbers

$\xi_1,\xi_2,\cdots\xi$ i.i.d with μ,σ^2, $E[\xi^4]<\infty$, Let $X_n=\bar{\xi}_n$ be the sample mean, then $P\{\lim_{n\to\infty}X_n=\mu\}=1$ (切比雪夫[illegible])

X_n converges in distribution ($X_n\xrightarrow{d}X$)

$\boxed{X_n\xrightarrow{d}X}$ if $\lim_{n\to\infty}F_n(x)=F(x)$ for all x at F(x) continuous

If $X_n\xrightarrow{P}X$, then $X_n\xrightarrow{d}X$

$X_n\xrightarrow{P}a\Longleftrightarrow X_n\xrightarrow{d}a$ (for constant a)

$\boxed{CLT}$: i.i.d random variables

Let $\xi_1,\xi_2,\cdots\xi$ i.i.d with $\mu=E[\xi]$ and $\sigma^2=Var(\xi)$

let $Z_n=\frac{\sqrt{n}(\bar{\xi}_n-\mu)}{\sigma}$. then $Z_n\sim N(0,1)$

$\boxed{\text{Normal approximation}}$

Let $S_n=\xi_1+\xi_2+\cdots+\xi_n$, $P\{S_n\leq s\}$:

(i) calc mean: $n\mu$ and var $n\sigma^2$ of S_n

(ii) calc $z=\frac{s-n\mu}{\sqrt{n}\sigma}$ (iii) $P\{S_n\leq s\}\approx\Phi(z)$

De Morivre-Laplace approximation.

if $S_n\sim Binomial(n,p)$, n is large and $k\leq l\geq0$

then $P\{k\leq S_n\leq l\}\approx\Phi\left(\frac{l+\frac{1}{2}-np}{\sqrt{np(1-p)}}\right)-\Phi\left(\frac{k-\frac{1}{2}-np}{\sqrt{np(1-p)}}\right)$

weak convergence

$\boxed{F_n\xrightarrow{w}F}$: $\lim_{n\to\infty}F_n(x)=F(x)$ (x continuous) $F_n\xrightarrow{w}F\Rightarrow X_n\xrightarrow{d}X$

$F_n\xrightarrow{w}F\Rightarrow X_n\xrightarrow{a.s.}X$ if X_n, X on $(\Omega,\mathcal{F},P)$, $X_n\sim F_n$, $X\sim F$

if $X_n\xrightarrow{a.s.}X$, let g be continuous, then $g(X_n)\xrightarrow{a.s.}g(X)$

if ..., for 连续.有界的g, $E[g(X_n)]\to E[g(X)]$

(i) $F_n\xrightarrow{w}F$ (ii) $X_n\xrightarrow{d}X$ (iii) $E[g(X_n)]\to E[g(X)]$ (三者等价)

$X_n\xrightarrow{d}X\Longleftrightarrow\lim_{n\to\infty}p_n(k)=p(k)$ for all $k\in N_0$

$X_n\xrightarrow{d}X\Rightarrow h(X_n)\xrightarrow{d}h(X)$ (h continuous)

g,g',g'',g''' 连续有界, then $E[g(X_n)]\to E[g(X)]\Rightarrow X_n\xrightarrow{d}X$

$F_n\xrightarrow{w}F\Rightarrow\varphi_n(t)\to\varphi(t)$ pointwise

$\varphi_n(t)=E[e^{itX_n}]$, $\varphi(t)=E[e^{itX}]$

$\varphi_n(t)\to\varphi(t)$, $t=0$ 连续 $\Rightarrow F_n\xrightarrow{w}F$

$X_n\xrightarrow{d}X$, $Y_n\xrightarrow{d}c$, then $X_n+Y_n\xrightarrow{d}X+c$, $X_nY_n\xrightarrow{d}cX$

$\boxed{X_n\xrightarrow{L^p}X}$: $\lim_{n\to\infty}E[|X_n-X|^p]=0$

在某些条件下强于P收敛

Bernoulli: $E[X]=p$, $Var(X)=p(1-p)$

Binom: $p(k)=\binom{n}{k}p^k(1-p)^{n-k}$ $E[X]=np$, $Var(X)=np(1-p)$

Geometric: $p(k)=(1-p)^{k-1}p$, $P\{X\geq k\}=(1-p)^{k-1}$

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

Poisson(λ): p很小, n很大, $p(k)=e^{-\lambda}\frac{\lambda^k}{k!}$

$E[X]=\lambda$, $Var(X)=\lambda$, Binom

Uniform(a,b): $f(x)=\begin{cases}\frac{1}{b-a}, & a\leq x\leq b\\ 0, & \text{otherwise}\end{cases}$ $E[X]=\frac{a+b}{2}$, $Var(X)=\frac{(b-a)^2}{12}$

Normal(μ,σ^2): $f(x)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$, $-\infty<x<\infty$

$E[X]=\mu$, $Var(X)=\sigma^2$

Exponential(λ): $f(x)=\begin{cases}\lambda e^{-\lambda x} & x\geq0\\ 0 & x<0\end{cases}$ 无记忆性

$F(x)=\begin{cases}1-e^{-\lambda x}, & x\geq0\\ 0, & x<0\end{cases}$ $E[X]=\frac{1}{\lambda}$, $Var(X)=\frac{1}{\lambda^2}$

Gamma(α,λ):

$f(x)=\begin{cases}\frac{\lambda e^{-\lambda x}(\lambda x)^{\alpha-1}}{\Gamma(\alpha)}=\frac{\lambda^\alpha}{\Gamma(\alpha)}e^{-\lambda x}x^{\alpha-1}, & x\geq0\\ 0, & x<0\end{cases}$

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

$\Gamma(\alpha)=(\alpha-1)\Gamma(\alpha-1)$ if $\alpha=n$, $\Gamma(n)=(n-1)!$

$E[X]=\frac{\alpha}{\lambda}$, $Var(X)=\frac{\alpha}{\lambda^2}$

Cauchy: $f(x)=\frac{1}{\pi}\frac{1}{1+x^2}$, $-\infty<x<\infty$

Beta: $f(x)=\begin{cases}\frac{1}{B(a,b)}x^{a-1}(1-x)^{b-1}, & 0<x<1\\ 0, & \text{otherwise}\end{cases}$

$B(a,b)=\int_0^1x^{a-1}(1-x)^{b-1}dx=\frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$

$E[X]=\frac{a}{a+b}$ $Var(X)=\frac{ab}{(a+b)^2(a+b+1)}$

Normal approximation to binom:

$P\{a\leq\frac{S_n-np}{\sqrt{np(1-p)}}\leq b\}\to\Phi(b)-\Phi(a)$

Jensen's inequality: $\varphi(E[X])\leq E[\varphi(X)]$ if φ is convex

$\int\tan u\,du=\ln|\sec u|+C$ $\int\sec u\,du=\ln|\sec u+\tan u|+C$

$\int\cot u\,du=\ln|\sin u|+C$ $\int\csc u\,du=-\ln|\csc u+\cot u|+C$

$\frac{d}{dx}(\tan x)=\sec^2x$ $\frac{d}{dx}(\cot x)=-\csc^2x$

$\frac{d}{dx}(\sec x)=\sec x\tan x$ $\frac{d}{dx}(\csc x)=-\csc x\cot x$

$\lim_{n\to\infty}(1+\frac{\lambda}{n})^n=e^\lambda$, $\sum_{i=0}^\infty\frac{x^i}{i!}=e^x$