FORMULA SHEET - PROBABILITY & STATISTICS

Euler's Gamma Function $\to \Gamma:(0,\infty)\to(0,\infty);\ \Gamma(1)=1$

$\boxed{\Gamma(a)=\int_0^\infty x^{a-1}e^{-x}\,dx}$ | $\Gamma(a+1)=a\Gamma(a)$, $\forall a>0$ | $\Gamma(m+1)=m!$, $\forall m\in\mathbb{N}$

$\Gamma\left(\frac{1}{2}\right)=\sqrt{2}\int_0^\infty e^{-\frac{t^2}{2}}\,dt=\sqrt{\pi}$

Euler's Beta Function $\to \beta:(0,\infty)\times(0,\infty)\to(0,\infty);\ \beta(a,1)=\frac{1}{a}$

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

$A_k^m=\frac{m!}{(m-k)!}$ | $P_m=m!=A_m^m$ | $C_m^k=\frac{A_m^k}{P_k}=\frac{m!}{k!(m-k)!}$

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

$P(A\setminus B)=P(A)-P(A\cap B)$ | $P(A|B)=\frac{P(A\cap B)}{P(B)}$, $P(B)\neq 0$ ← condition

A,B - indep $\Leftrightarrow P(A\cap B)=P(A)P(B)\Leftrightarrow P(A|B)=P(A)$

$\{A_i\}_{i\in I}$ partition of $S\Rightarrow P(E)=\sum_{i\in I}P(A_i)P(E|A_i)$

BINOMIAL MODEL
- k successes in n Bernoulli trials, with p - probability of success ($q=1-p$)

$\boxed{P(n,k)=C_n^k p^k q^{n-k},\ k=\overline{0,n}}$

- each trial has two possible outcomes: Fail / Succ
- the prob. for success is the same for each trial; the parameters: n - no. trials, p - prob of S.
- $P(n,k)$ is the coef of x^k in the Bino Model

$(px+q)^n=\sum_{k=0}^n C_n^k p^k q^{n-k}x^k$

HYPERGEOMETRIC MODEL
- Binomial without replacement (some objects in the sample have a certain property; parameters:

N - total no. of obj [redacted], n_1 - obj. with the prop.; n - no. of trials (k - NOT a param, ever)

$\boxed{P(n;k)=\frac{C_{n_1}^k C_{N-n_1}^{n-k}}{C_N^n}}$ $\sum_{k=0}^n P(n,k)=1$

POISSON MODEL
- prob of K successes in n trials, each trial has its own prob of succ.

$\boxed{P(n;k)=\sum_{1\le i_1<\dots<i_k\le n} p_{i_1}\dots p_{i_k}q_{i_{k+1}}\dots q_n}$

$P(n,k)$ - coef of x^k in the polynom. exp.

$(p_1x+q_1)(p_2x+q_2)\dots(p_nx+q_n)$

PASCAL (NEGATIVE BINOMIAL) MODEL
↳ the probability of n-th success after k failures in a seq. of Bernoulli trials with prob of success $=p$

$\boxed{P(n;k)=C_{n+k-1}^{n-1}p^nq^k=C_{n+k-1}^{k}p^nq^k}$

$P(n,k)$ is the coef of x^k in

$\left(\frac{p}{1-qx}\right)^n=\sum_{k=0}^\infty P(n,k)x^k,\ |qx|<1$

GEOMETRIC MODEL
- Pascal, $n=1$

↳ 1st success after k failures in a sequence of Bernoulli trials

$\boxed{p_k=pq^k}$, p_k is the coef of x_k

$\frac{p}{1-qx}=\sum_{k=0}^\infty p_kx^k,\ |qx|<1$

↳ if one wants the no. of trials before the first success ($Y=X+1$) $\Rightarrow$

$\tilde{p}_k=pq^{k-1}$

BERNOULLI DISTRIBUTION $\boxed{Bern(p)}$
pdf: $X\begin{pmatrix}0 & 1\\ 1-p & p\end{pmatrix}$, $p\in(0,1)$

BINOMIAL DISTRIBUTION $Bino(n,p)$
pdf: $X\begin{pmatrix}k\\ C_n^kp^kq^{n-k}\end{pmatrix}_{k=\overline{0,n}}$, $p\in(0,1)$, $n\in\mathbb{N}$ $\boxed{X\in B(n,p)}$

DISCRETE UNIFORM DIST. $\boxed{U(m)}$
pdf $X\begin{pmatrix}k\\ \frac{1}{m}\end{pmatrix}_{k=\overline{1,m}}$, $m\in\mathbb{N}$

HGEOMETRIC DIST. $\boxed{H(N,n_1,n)}$
pdf: $X\begin{pmatrix}k\\ \frac{C_{n_1}^kC_{N-n_1}^{n-k}}{C_N^n}\end{pmatrix}_{k=\overline{0,n}}$, $N,n_1,n\in\mathbb{N}$

POISSON DIST. $\boxed{P(\lambda)}$ $\lambda>0$
$X\begin{pmatrix}k\\ \frac{\lambda^k}{k!}e^{-\lambda}\end{pmatrix}_{k=0,1,\dots}$

- law of rare events, not based on the poisson model;
- involves observing discrete events over a cont. period of TIME, length, space, etc.

NEGATIVE BINO. (PASCAL)
$\boxed{NB(n,p)}$, $n\in\mathbb{N}$, $p\in(0,1)$

$X\begin{pmatrix}k\\ C_{n+k-1}^k p^n q^k\end{pmatrix}_{k=0,1,\dots}$

GEOMETRIC $p\in(0,1)$ $\boxed{Geo(p)}$
$X\begin{pmatrix}k\\ pq^k\end{pmatrix}_{k=0,1,\dots}$

CUMULATIVE DIST. FUNC. (CDF)
$F_X:\mathbb{R}\to\mathbb{R},\ F_X(x)=P(X\le x)=\sum_{x_i\le x}p_i$

$(X,Y):S\to\mathbb{R}^2$ - disc. rand vec.

- joint pdf $p_{ij}=P(X=x,Y=y)$, $i,j\in I\times J$
- joint cdf $F=F(x,y):\mathbb{R}\to\mathbb{R}$, $F(X\le x,Y\le y)=\sum_{x_i\le x}\sum_{y_j\le y}p_{ij}$, $\forall(x,y)\in\mathbb{R}^2$

MARGINAL DENSITIES: $p_i=P(X=x_i)=\sum_{j\in J}p_{ij}$; $q_j=P(Y=y_j)=\sum_{i\in I}p_{ij}$; $\boxed{!!}$ X,Y indep $\Rightarrow$ $P(X=x_i,Y=y_j)=P(X=x_i)P(Y=y_j)=p_iq_j$

$X+Y\begin{pmatrix}x_i+y_j\\ p_{ij}\end{pmatrix}$ $\alpha X\begin{pmatrix}\alpha x_i\\ p_i\end{pmatrix}$ $XY\begin{pmatrix}x_iy_j\\ p_{ij}\end{pmatrix}$ $X/Y\begin{pmatrix}x_i/y_j\\ p_{ij}\end{pmatrix}$

X - cont rand var, pdf f

1) $F(x)=P(X\le x)=\int_{-\infty}^x f(t)\,dt$
2) $f(x)\ge 0,\ \forall x\in\mathbb{R}$, $\int_{\mathbb{R}}f(x)\,dx=1$
3) $P(X=x)=0$; $P(a<X<b)=P(a\le X\le b)=\int_a^b f(t)\,dt$
4) $F(-\infty)=0,\ F(\infty)=1$

$(X,Y):S\to\mathbb{R}^2$, pdf $f(x,y):\mathbb{R}^2\to\mathbb{R}^2$, cdf $F_{(x,y)}$

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

1. $P(a_1<X\le b_1,a_2<Y\le b_2)=F(b_1,b_2)-F(a_1,b_2)-F(b_1,a_2)+F(a_1,a_2)$
2. $F(\infty,\infty)=1,\ F(-\infty,y)=F(x,-\infty)=0$
3. $P((X,Y)\in D)=\iint_D f(x,y)\,dy\,dx$

$f_X(x)=\int_{\mathbb{R}}f(x,y)\,dy$, $f_Y(y)=\int_{\mathbb{R}}f(x,y)\,dx$ } marginal densities

X,Y - indep $\Rightarrow f_{(X,Y)}(x,y)=f_X(x)f_Y(y),\ \forall(x,y)\in\mathbb{R}^2$

$Y=g(X)$, $g'\neq 0$, monot. $g:\mathbb{R}\to\mathbb{R}$

$f_Y(y)=\frac{f_X(g^{-1}(y))}{g'(g^{-1}(y))}$, $y\in g(\mathbb{R})$! Important

pdf Y

UNIFORM DIST. $U(a,b)$ $-\infty<a<b<\infty$
pdf $f(x)=\frac{1}{b-a}$, $x\in[a,b]$

NORMAL DIST $N(\mu,\sigma)$, $\sigma>0$
pdf $f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$, $x\in\mathbb{R}$

GAMMA DIST $Gamma(a,b)$ $a,b>0$
pdf: $f(x)=\frac{1}{\Gamma(a)b^a}x^{a-1}e^{-\frac{x}{b}}$, $x>0$

↳ models the total time of a multistage sc[enario]

EXPONENTIAL $EXP(\lambda)=Gamma(1,1/\lambda)$, $\lambda>[0]$
↳ models time: waiting time, intrarrival, failure, time between rare events, etc.

↳ $\alpha\in\mathbb{N}$, $Gamma(\alpha,1/\lambda)$ var $=\sum_{i=1}^\alpha Exp(\lambda)$ vars

Expectation - X discrete $\Rightarrow E(X)=\sum_{i\in I}x_ip_i$

↳ X continuous, with pdf $f\Rightarrow E(X)=\int_{\mathbb{R}}xf(x)\,dx$

Variance $V(X)=E(X^2)-(E(X))^2=E((X-E(X))^2)$

Std Deviation $\sigma(X)=Std(X)=\sqrt{V(X)}$

Moments - of order k: $\nu_k=E(X^k)$

↳ abs. mom. of order k: $\underline{\nu}_k=E(|X|^k)$

↳ central mom of order k: $\mu_k=E((X-E(X))^k)$

I. $E(aX+b)=aE(X)+b$; $V(aX+b)=a^2V(X)$

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

III. X,Y indep: $E(XY)=E(X)E(Y)$ $V(X\pm Y)=V(X)+V(Y)$

IV. $E(h(X))=\sum h(x_i)p_i$ (disc) $E(h(X))=\int_{\mathbb{R}}h(x)f(x)\,dx$

COVARIANCE $cov(X,Y)=E((X-E(X))(Y-E(Y)))$

CORRELATION COEF. $\rho(X,Y)=\frac{cov(X,Y)}{\sqrt{V(X)}\sqrt{V(Y)}}$; $cov(X,Y)=E(XY)-E(X)$

$V\left(\sum_{i=1}^n a_iX_i\right)=\sum_{i=1}^n a_i^2V(X)+2$

$E(h(X,Y))=\int_{-\infty}^\infty\int_{-\infty}^\infty h(x,y)f(x,y)\,dx\,dy$

CENTRAL LIMIT THEOREM (CLT) $X_1\dots X_n$ indep

$\mu=E(X_i)$; $\sigma=\sigma(X_i)=Std(X_i)$; $S_n=\sum_{i=1}^n X_i$; n —

$Z_n=\frac{S_n-E(S_n)}{\sigma(S_n)}=\frac{S_n-n\mu}{\sigma\sqrt{n}}\to Z\in N(0,1)$; $F_{Z_n}\xrightarrow{cdf}F_Z$

$\Phi=F_Z(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^x e^{-\frac{t^2}{2}}\,dt$ LAPLACE

CHEBYSHEV $P(|X-E(X)|\ge\varepsilon)\le\frac{1}{\varepsilon^2}V(X)$, $\forall\varepsilon>0$

$P(|X-E(X)|<\varepsilon)\ge 1-\frac{1}{\varepsilon^2}V(X)$

MARKOV $P(|X|\ge a)\le\frac{1}{a}E(|X|)$

LYAPUNOV $(E(|X-c|^a))^{\frac{1}{a}}\le(E(|X-c|^b))^{\frac{1}{b}}$,

CAUCHY-B. $E(|X|)\le\sqrt{E(X^2)}$

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

Method of Moments

$v_k = \bar{v}_k$, $v_k = E(x)$ cont

$\bar{v}_k$ - mean $= \sum_{i=1}^{m} x_i \cdot \frac{1}{m}$

Method of Max. Likelyhood

$\frac{\partial \ln L(x_1 \dots x_m; \theta)}{\partial \theta_j} = 0,$

$L(x_1 \dots x_m; \theta) = \prod_{j=1}^{m} f(x_j; \theta)$

Std Error of an Estimator

$\sigma_{\hat{\theta}} = \sigma(\bar{\theta}) = \sqrt{V(\bar{\theta})}$

FISHER INFORMATION

$I_m(\theta) = -E\left[\frac{\partial^2 \ln L(X; \theta)}{\partial \theta^2}\right]$

° range of X indep from $\theta \Rightarrow I_m(\theta) = m I_1(\theta)$

ABSOLUTELY CORRECT EST.

I. $E(\hat{\theta}) = \theta$ (unbiased)

II. $\lim_{m \to \infty} V(\bar{\theta}) = 0$

EFFICIENCY OF AN ABS CORRECT EST.

$e(\bar{\theta}) = \frac{1}{I_m(\theta) V(\bar{\theta})}$

efficient if $e(\bar{\theta}) = 1$

MVUE: $E(\bar{\theta}) = 0$, $V(\bar{\theta}) < V(\theta)$

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

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

THE PROBABILITY RULE

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

$N(0,1)$ denoted by Z

$f_Z(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$

Variance

$V(X) = \sum_{i \in I} x_i^2 p_i - \left(\sum_{i \in I} x_i p_i\right)^2$

$V(X) = \int_{\mathbb{R}} x^2 f(x) dx - \left(\int_{\mathbb{R}} x f(x) dx\right)^2$

$f_Y(y) = \frac{f_X(g^{-1}(y))}{|g'(g^{-1}(y))|}$

The power of a test on a param θ = prob of rejecting the null hypothesis

$\pi(\theta^*) = P(\text{reject } H_0 | \theta = \theta^*) = P(TS \in RR | \theta = \theta^*) = 1 - P(\text{not reject } H_0 | H_1) = 1 - \beta(\theta_1)$

We preset $\alpha = \pi(\theta_0)$ we det a RR for which $\pi(\theta_1) = 1 - \beta(\theta_1)$ the largest possible $\Rightarrow$ most powerful test

NEYMAN-PEARSON LEMMA

$H_0: \theta = \theta_0$

$H_1: \theta = \theta_1$

$L(\theta) = L(X_1, \dots, X_m)$

For a fixed $\alpha \in (0,1)$, a most powerful test is the test with its $RR = \left\{ \frac{L(\theta_1)}{L(\theta_0)} \geq k_\alpha \right\}$

↳ write X in terms of k_α

$\alpha = P(X_1 \in RR | H_0) = P(X_1 \geq k_\alpha | \theta = 1)$

The power of a test:

$\pi(\theta_1) = 1 - \beta(\theta_1)$