Formula Sheet

Counting and Basic Probability

$$P_r^n = n(n-1)\cdots(n-r+1) = \frac{n!}{(n-r)!} \qquad C_r^n = \binom{n}{r} = \frac{n!}{r!(n-r)!} \qquad \binom{n}{n_1, n_2, \cdots, n_k} = \frac{n!}{n_1! n_2! \cdots n_k!}$$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \qquad P(A) + P(A^C) = 1 \qquad P(A \cap B) = P(A|B)P(B) \qquad P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$P(A) = \sum_{i=1}^{k} P(B_i)P(A|B_i) \qquad P(B_j|A) = \frac{P(B_j)P(A|B_j)}{P(A)} = \frac{P(B_j)P(A|B_j)}{\sum_{i=1}^{k} P(B_i)P(A|B_i)}$$

Random Variable Definitions and such

$$p_Y(y) = P(Y = y) \qquad F_Y(y) = P(Y \leq y) \qquad f_Y(y) = \frac{d}{dy} F_Y(y) \qquad F_Y(y) = \int_{-\infty}^{y} f_Y(t)dt$$

pmf/pdf must have two properties: $f_Y(y) \geq 0$ or $p_Y(y) \geq 0$ and integral or sum over entire support must be 1.

$$E(g(Y)) = \begin{cases} \sum_y g(y)p_Y(y) & \text{if Y is discrete} \\ \int_{-\infty}^{\infty} g(y)f_Y(y)dy & \text{if Y is continuous} \end{cases} \qquad P(a \leq Y \leq b) = \begin{cases} \sum_{y: a \leq y \leq b} p_Y(y) & \text{if Y is discrete} \\ \int_a^b f_Y(y)dy & \text{if Y is continuous} \end{cases}$$

$$Mean: \ \mu_Y = E(Y) \qquad Variance: \ \sigma_Y^2 = Var(Y) = E\left[(Y - \mu_Y)^2\right] = E(Y^2) - \mu_Y^2$$

$$m_Y(t) = E(e^{tY}) \qquad E(Y^n) = m_Y^{(n)}(0) = \frac{d^n}{dt^n} m_Y(t)|_{t=0}$$

$E(a + bY) = a + bE(Y)$ $\quad Var(a + bY) = b^2 Var(Y)$ $\quad p^{th}$ quantile of $F_Y(y)$ is value y_p s.t $F_Y(y_p) = p$

Random Math Things

Binomial Theorem: $(x+z)^n = \sum_{y=0}^{n} \binom{n}{y} x^y z^{n-y}$ $\qquad$ Maclaurin Series of $e^\lambda = \sum_{y=0}^{\infty} \frac{\lambda^y}{y!}$

Finite Sum of Geometric: $\sum_{y=0}^{n-1} ap^y = \frac{a(1-p^n)}{1-p}$ $\qquad$ Sum of Geometric: $\sum_{y=0}^{\infty} ap^y = \frac{a}{1-p}$

Other geometric sum properties: $\sum_{y=1}^{\infty} yp^{y-1} = \frac{1}{(1-p)^2}$ $\qquad \sum_{y=0}^{\infty} \frac{p^{y+1}}{y+1} = -ln(1-p)$

If $a_1, \ldots, a_n$ converges to a then $\left(1 + \frac{a_n}{n}\right)^n \to e^a$

$$\Gamma(\alpha) = \int_0^{\infty} t^{\alpha-1} e^{-t} dt \qquad \Gamma(\alpha+1) = \alpha\Gamma(\alpha) \ for \ \alpha > 0 \qquad \Gamma(n+1) = n! \ for \ n \ integer \qquad \Gamma(1/2) = \sqrt{\pi}$$

Discrete Distributions

Binomial: $p_Y(y)=\binom{n}{y}p^y(1-p)^{n-y}$ *support*: $y=0,1,2,....,n$

$$E(Y)=np \quad Var(Y)=np(1-p) \quad M_Y(t)=(pe^t+(1-p))^n$$

Geometric: $p_Y(y)=(1-p)^{y-1}p$ *support* $y=1,2,....$ For Integer $b, P(Y>b)=(1-p)^b$

$$E(Y)=1/p \quad Var(Y)=(1-p)/p^2 \quad M_Y(t)=\frac{pe^t}{1-(1-p)e^t} \quad t<-ln(1-p)$$

Neg Bin: $p_Y(y)=\binom{y-1}{r-1}p^r(1-p)^{y-r}$ *support*: $y=r,r+1,r+2....$

$$E(Y)=r/p \quad Var(Y)=r\frac{(1-p)}{p^2} \quad M_Y(t)=\left(\frac{e^tp}{1-(1-p)e^t}\right)^r \quad t<-ln(1-p)$$

Hypergeometric: $p_Y(y)=\frac{\binom{r}{y}\binom{n-r}{m-y}}{\binom{n}{m}}$ *support*: $y=0,1,2,...m$

$$E(Y)=rm/n \quad Var(Y)=\frac{mr(n-r)(n-m)}{n^2(n-1)} \quad No\ closed\ form\ MGF$$

Poisson: $p_Y(y)=\frac{\lambda^y e^{-\lambda}}{y!}$ *support*: $y=0,1,2,....$

$$E(Y)=\lambda \quad Var(Y)=\lambda \quad M_Y(t)=e^{\lambda(e^t-1)}$$

Continuous Distributions

Gamma: $f_Y(y)=\frac{\lambda^\alpha}{\Gamma(\alpha)}y^{\alpha-1}e^{-\lambda y}$ *support*: $0<y<\infty$

$$E(Y)=\alpha/\lambda \quad Var(Y)=\alpha/\lambda^2 \quad M_Y(t)=\left(\frac{1}{1-t/\lambda}\right)^\alpha, \ t<\lambda$$

If $\alpha=1$ then $Y\sim exp(\lambda)$. If $\alpha=p/2$, $\lambda=1/2$, then $Y\sim\chi^2_p$

Normal: $f_Y(y)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2\sigma^2}(y-\mu)^2}$ *support*: $-\infty<y<\infty$

$$E(Y)=\mu \quad Var(Y)=\sigma^2 \quad M_Y(t)=e^{\mu t+\sigma^2t^2/2}$$

If $Z=\frac{Y-\mu}{\sigma}$ then $Z\sim N(0,1)$.

Beta: $f_Y(y)=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}y^{\alpha-1}(1-y)^{\beta-1}$ *support*: $y\in[0,1]$

$$E(Y)=\frac{\alpha}{\alpha+\beta} \quad Var(Y)=\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} \quad M_Y(t)=Not\ useful$$

Uniform: $f_Y(y)=\frac{1}{\theta_2-\theta_1}$ *support*: $y\in[\theta_1,\theta_2]$

$$E(Y)=\frac{\theta_1+\theta_2}{2} \quad Var(Y)=\frac{(\theta_2-\theta_1)^2}{12} \quad M_Y(t)=\frac{e^{\theta_2t}-e^{\theta_1t}}{(\theta_2-\theta_1)t}$$

Transformation Theorems

If $X \sim N(\mu, \sigma^2)$ then $Y = aX + b \sim N(a\mu + b, a^2\sigma^2)$

If $X \sim f_X(x)$, a continuous RV, and $Y = g(X)$, where g is differentiable and strictly monotonic over the support of X, then $f_Y(y) = f_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right|$

Probability Integral Transform: $X \sim F_X(x)$, a continuous RV with strictly increasing cdf, $Y = F_X(x) \sim U(0,1)$

Let $U \sim U(0,1)$ and define $X = F^{-1}(U)$, then X has cdf F.

Joint Distributions

$$F_{X,Y}(x,y) = P(X \le x, Y \le y) \quad p_{X,Y}(x,y) = P(X = x, Y = y) \quad p_X(x) = \sum_y p_{X,Y}(x,y) \quad P((X,Y) \in A) = \sum_{(x,y)\in A} p_{X,Y}(x,y)$$

$Multinomial: p_{Y_1,\dots,Y_n}(y_1,\dots,y_n|p's) = \binom{m}{y_1 \dots y_n} \Pi_{i=1}^n p_i^{y_i} \quad support: y_i \ge 0,\ y_1 + \dots + y_n = m$

Marginally, $Y_i \sim Bin(m, p_i)$

$$F_{X,Y}(x,y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f_{X,Y}(s,t)\,ds\,dt \quad \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x,y) = f_{X,Y}(x,y) \quad P(a \le X \le b, c \le Y \le d) = \int_a^b \int_c^d f_{X,Y}(x,y)\,dy\,dx$$

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

If (X,Y) $BVN(\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2, \rho)$ then marginally, $X \sim N(\mu_X, \sigma_X^2)$ and $Y \sim N(\mu_Y, \sigma_Y^2)$ and conditionally, $X|Y = y \sim N(\mu_X + \rho \frac{\sigma_X}{\sigma_Y}(y - \mu_Y), \sigma_X^2(1 - \rho^2))$

Conditional Distributions

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

If joint cdf/pdf/pmf is product of marginal cdf/pdf/pmf then $Y_1, Y_2, \dots, Y_n$ are independent.

X is independent of Y if and only if $f_{X,Y}(x,y) = g(x)h(y)$

Bayesian Posterior Distribution: $f_{\Theta|Y}(\theta|y) = \frac{f_{Y|\Theta}(y|\theta) f_\Theta(\theta)}{f_Y(y)} \propto f_{Y|\Theta}(y|\theta) f_\Theta(\theta)$

Order Statistics

Distribution of j^{th} order statistic: $f_{Y_{(j)}}(y) = n\binom{n-1}{j-1} f_Y(y)[F_Y(y)]^{j-1}[1 - F_Y(y)]^{n-j}$

Joint Distribution of i^{th} and j^{th} order statistic:

$$f_{Y_{(i)},Y_{(j)}}(u,v) = n f_Y(u)(n-1) f_Y(v) \binom{n-2}{i-1, j-i-1, n-j} [F_Y(u)]^{i-1}[F_Y(v) - F_Y(u)]^{j-i-1}[1 - F_Y(v)]^{n-j}$$

Expected Values of a function of multiple RVs

$$E(g(Y_1,...,Y_n)) = \begin{cases} \sum_{y_1}...\sum_{y_n} g(y_1,...,y_n) p_{Y_1,...,Y_n}(y_1,...,y_n) \\ \int_{y_1}...\int_{y_n} g(y_1,...,y_n) f_{Y_1,...,Y_n}(y_1,...,y_n) dy_n...dy_1 \end{cases}$$

If X is independent of Y then $E(g(X)h(Y)) = E(g(X))E(h(Y))$

$$Cov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)] = E(XY) - E(X)E(Y) \quad Corr(X,Y) = \rho_{XY} = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$$

$$Var\left(a + \sum_{i=1}^{n} b_i X_i\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} b_i b_j Cov(X_i,X_j) = \sum_{i=1}^{n} b_i^2 Var(X_i) + 2\sum_{i<j} b_i b_j Cov(X_i,X_j)$$

$$E(g(X)|Y=y)) = \begin{cases} \sum_{x|y} g(x) p_{X|Y}(x|y) \\ \int_{x|y} g(x) f_{X|Y}(x|y) dx \end{cases} \qquad E(X) = E(E(X|Y)) \quad Var(X) = E(Var(X|Y)) + Var(E(X|Y))$$

Delta Method approx for $Y = g(X)$, (first order) $\mu_Y \approx g(\mu_X)$, $\sigma_Y^2 \approx [g'(\mu_x)]^2 \sigma_X^2$ (second order) $\mu_Y \approx g(\mu_X) + \frac{1}{2} g''(\mu_X) \sigma_X^2$

Limit Theorems etc.

$$Y_n \xrightarrow{p} Y \quad if \lim_{n\to\infty} P(|Y_n - Y| \geq \epsilon) = 0 \quad or \lim_{n\to\infty} P(|Y_n - Y| < \epsilon) = 1$$

$$Y_n \xrightarrow{d} Y \quad if \lim_{n\to\infty} F_{Y_n}(y) = F_Y(y) \quad or \lim_{n\to\infty} M_{Y_n}(y) = M_Y(y)$$

Convergence in probability implies convergence in distribution. The converse is true when you converge in distribution to a constant.

Markov's Inequality - If X is a nonnegative RV where $E|X| < \infty$ then for $t > 0$, $P(X \geq t) \leq E(X)/t$

Chebychev's Inequality - X a RV with μ, σ^2, then for $t > 0$ $P(|X - \mu| \geq t) \leq \sigma^2/t^2$

WLLN - Y_n a sequence of independent RVs with $E(Y_i) = \mu$ then $\bar{Y_n} = \sum_{i=1}^{n} Y_i/n \xrightarrow{p} \mu$

Continuity Theorem - If Y_n converges to Y, g a continuous ftn, then $g(Y_n)$ converges to $g(Y)$

CLT - Y_i iid with $\sigma^2 < \infty$ then $\frac{\bar{Y}_n - \mu}{\sigma/\sqrt{n}} \xrightarrow{d} N(0,1)$ *or* $\bar{Y}_n \dot{\sim} N(\mu, \sigma^2/n)$ *or* $\sqrt{n}(\bar{Y}_n - \mu) \xrightarrow{d} N(0,\sigma^2)$

Delta method normality - Y_n RVs where $Y_n \dot{\sim} N(\theta_0, \sigma^2)$ then for a function g and value θ_0 where $g'(\theta_0)$ exists and is not 0 $g(Y_n) \dot{\sim} N\big(g(\theta_0), [g'(\theta_0)]^2 \sigma^2\big)$.

Monte Carlo Integration - $(b-a)\frac{1}{n}\sum_{i=1}^{n} g(X_i) \xrightarrow{p} \int_a^b g(x)dx$

Slutsky's Theorem - If $X_n \xrightarrow{d} X, Y_n \xrightarrow{p} a$ then $X_n Y_n \xrightarrow{d} aX$ and $X_n + Y_n \xrightarrow{d} X + a$

Estimators - generic estimator notation $\hat{\underset{\sim}{\theta}}$, truth θ

Bias - $Bias\left(\hat{\underset{\sim}{\theta}}\right) = E\left(\hat{\underset{\sim}{\theta}} - \theta\right)$

Variance - $Var\left(\hat{\underset{\sim}{\theta}}\right) = E\left[\left(\hat{\underset{\sim}{\theta}} - E\left(\hat{\underset{\sim}{\theta}}\right)\right)^2\right] = E\left(\hat{\underset{\sim}{\theta}}^2\right) - \left[E\left(\hat{\underset{\sim}{\theta}}\right)\right]^2$

MSE - Mean squared error = $MSE\left(\hat{\underset{\sim}{\theta}}\right) = E\left[\left(\hat{\underset{\sim}{\theta}} - \theta\right)^2\right] = \left[Bias\left(\hat{\underset{\sim}{\theta}}\right)\right]^2 + Var\left(\hat{\underset{\sim}{\theta}}\right)$