Probability Exam Sheet by David A Wheeler

GMU, MATH 351, Spring 1986, Dr. Bolstein, for textbook "Introductory Probability and Statistical Applications," 2nd edition by Paul L. Meyer

$\int u^n du = \frac{u^{n+1}}{n+1} + C,\ n \neq -1$

Ch 1

$a \in B$ (Def) a is member of B — $A \subset B$ (Def) A is a subset of B — $\int \frac{du}{u} = \ln|u| + C$

$A \cup B$ A or B — $A \cap B$ A and B — $\int \cos u\, du = \sin u + C$

$uv = \int u\,dv + \int v\,du$; $\int a^u du = \frac{1}{\ln a} a^u + C$; $\int \sin u\, du = -\cos u + C$; $\int \tan u\, du = \ln|\sec u| + C$

$\int \frac{du}{\sqrt{a^2-u^2}} = \sin^{-1}\frac{u}{a} + C$

$P(A) = 1 - P(\bar{A})$ — if $A \subset B$, $P(A) \leq P(B)$

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$\int \frac{du}{a^2+u^2} = \frac{1}{a}\tan^{-1}\frac{u}{a} + C$

$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$

if =ly likely outcomes $P(A) = \frac{\#\text{ possible favorable}}{\#\text{ possible}}$

$\int \frac{du}{u\sqrt{u^2-a^2}} = \frac{1}{a}\sec^{-1}\frac{u}{a} + C$

Factorial $n! = (n)(n-1)\cdots(1)$

Permutation: n different object, choose r and permute (order important)

$\int \sin u\, du = -\cos u + C$

$_nP_r = \frac{n!}{(n-r)!}$

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

Combination: order unimportant $\binom{n}{r} = \frac{n!}{r!\,(n-r)!} = \binom{n}{n-r} = \binom{n-1}{r-1} + \binom{n-1}{r}$

Hypergeometric: Given N objects, choose n of them. Probability of s_1 A's and s_2 B's, where $r_1 + r_2 = N$, total of r_1 A's and r_2 B's is

$$\frac{\binom{r_1}{s_1}\binom{r_2}{n-s_1}}{\binom{N}{n}}$$

$\frac{d}{dx}u^c = cu^{c-1}\frac{du}{dx}$

$\frac{d}{dx}uv = u\frac{dv}{dx} + v\frac{du}{dx}$

if not all different, but n_1 of 1, n_2 of 2, $n_1 + \cdots n_k = n$; Permutation $= \frac{n!}{n_1!\,n_2!\cdots n_k!}$

$\frac{d}{dx}\left(\frac{u}{v}\right) = \frac{v\frac{du}{dx} - u\frac{dv}{dx}}{v^2}$

$\frac{dy}{dx} = \frac{dy}{du}\cdot\frac{du}{dx}$ $y = f(u)$, $u = g(x)$

Multiplication Theorem

$P(B|A) = \frac{P(A \cap B)}{P(A)}$ Prob. of B, given A occurred. So $P(A \cap B) = P(A)P(B|A)$

$\frac{d}{dx}\sin u = \cos u\frac{du}{dx}$

$\frac{d}{dx}\tan u = \sec^2 u\frac{du}{dx}$

Baye's Theorem (find cause) $P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_{j=1}^{k} P(A|B_j)P(B_j)}$

where one & only one B_i can occur (partitioned)

Law of Total Prob: $P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \cdots$ if $B_{1,2\ldots}$ partition

Independant Events: iff $P(A \cap B) = P(A)P(B)$

A, B, C mutually independant iff: $P(A \cap B) = P(A)P(B)$, $P(A \cap C) = P(A)P(C)$, $P(B \cap C) = P(B)P(C)$, $P(A \cap B \cap C) = P(A)P(B)P(C)$

Sphere Surface Area $= 4\pi r^2$; Sphere Volume $= \frac{4}{3}\pi r^3$

$\cos x = \Sigma$

$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$ $\sin x = \sum_{i=1}^{\infty} (-1)^{i-1} \frac{x^{2i-1}}{(2i-1)!}$

Ch 4 Geometric Series $1 + r + r^2 + \cdots = \frac{1}{1-r}$ if $|r| < 1$.

BINOMIAL distribution: X be binomial variable

{given n $\boxed{\text{independant}}$ trials, prob p that event is <u>true</u>,

<u>x</u> = # of occurances of event in n trials}

$P[X=x] = \binom{n}{x} p^x q^{n-x}$ where $q = 1-p$

if x left as variable this is <u>probability</u> <u>mass</u> <u>function</u> (pmf)

<u>pmf</u>: largest value is at $\begin{cases} x = (n+1)p - 1 \text{ AND } (n+1)p & \text{if } (n+1)p \text{ integer} \\ x = INT((n+1)p) & \text{if } (n+1)p \text{ not integer} \end{cases}$

<u>pdf</u>: probability density function (for continuous functions): pdf of X

$f(x) \geq 0$ for all x, $\int_{-\infty}^{+\infty} f(x)dx = 1$, $P(a \leq X \leq b) = \int_a^b f(x)\,dx$

cdf: cumulative distribution function: $F(x) = P(X \leq x)$. If X discrete,

$F(x) = \sum_j p(x_j)$ where $x_j < x$; continuous, $F(x) = \int_{-\infty}^{x} f(s)ds$ and pdf $f(x) = F'(x)$ (cdf)

Ch5 if X cont. RV, pdf f, H is cont function, $Y = H(X)$, to find pdf $g(y)$: 1. find cdf of Y $G(y) = P(Y \leq y)$ by subst. Y 2. pdf $g(y) = \frac{dG(y)}{dy}$ 3. find where $f(x) \geq 0$.

(Ex) pdf $f(x) = \begin{cases} 2x & 0 \leq x \leq 1 \\ 0 & \text{else} \end{cases}$; $H(x) = 3x+1$; $G(y) = P(Y \leq y) = P(3X+1 \leq y) = \int_0^{(y-1)/3} 2x\,dx$; $g(y) = G'(y)$

Ch6 (X, Y): discrete: prob function $p(x_i, y_j) \geq 0$, $\sum\sum p(x_i, y_j) = 1$

continuous: joint prob. density $f(x,y) \geq 0$; $\int\int_R f(x,y)dx\,dy = 1$

cdf $F(x,y) = P(X \leq x, Y \leq y)$; $\partial^2 F(x,y)/\partial x \partial y = f(x,y)$

Marginal prob. functions: $g(x) = \int_{-\infty}^{\infty} f(x,y)dy$, $h(y) = \int_{-\infty}^{\infty} f(x,y)dx$

Uniformly distributed if $f(x,y) = \{\frac{1}{\text{area}}$ if $(x,y) \in R$, 0 else}

Conditional pdf of X for Y=y $g(x|y) = \frac{f(x,y)}{h(y)}$ $(h(y) > 0)$

X & Y independant iff $p(x_i, y_j) = p(x_i)p(y_j)$ always, $f(x,y) = g(x)h(y)$.

INDEPEND.

$\left\{\begin{array}{l} \text{if } W = XY, \text{ pdf } p(w) = \int_{-\infty}^{\infty} g(u)\, h\left(\frac{w}{u}\right) \left|\frac{1}{u}\right| du \\ \text{if } W = X/Y, \text{ pdf } q(z) = \int_{-\infty}^{\infty} g(vz)\, h(v)\, |v|\, dv \end{array}\right\}$ ← independ, continuous

if (X, Y) continuous, joint pdf f, let $Z = H_1(X,Y)$, $W = H_2(X,Y)$, then joint pdf of (Z, W) is $k(z,w) = f(G_1(z,w), G_2(z,w))\,|J(z,w)|$ where G_1 & G_2 are found by $x = G_1(z,w)$; $y = G_2(z,w)$; and $J(z,w) = \begin{vmatrix} \partial x/\partial z & \partial x/\partial w \\ \partial y/\partial z & \partial y/\partial w \end{vmatrix}$

Chebyshev's: $P[|X-\mu| \geq k\sigma] \leq k^{-2}$; $P[|X-c| \geq \epsilon] \leq \frac{1}{\epsilon^2} E[X-c]^2$

Ch 7) $E(X) = \begin{cases} \text{discrete } \sum x_i \underbrace{p(x_i)}_{\text{pmf}} = \sum H(y_i)p(y_i) & \text{if } X=H(Y) \\ \text{continuous } \int_{-\infty}^{\infty} x \underbrace{f(x)}_{\text{pdf}} dx = \int_{-\infty}^{\infty} H(y)f(y)dy \end{cases}$

If $Z=H(X,Y)$, $E(Z) = \begin{cases} \sum^{j}\sum^{i} H(x_i, y_j)\, p(x_i, y_j) \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H(x,y) f(x,y)\, dx\,dy \end{cases}$ joint pdf

if $X = C$ (constant), $E(X)=C$, $V(X)=0$; $E(CX) = CE(X)$; $E(X+Y)=E(X)+E(Y)$;

if independ, $E(XY) = E(X)E(Y)$; $V[\sum X_i] = \sum V[X_i] + 2\sum_{i<j}\sum Cov[X_i, X_j]$

Covariance $Cov(X,Y) = E[(X-\mu_x)(Y-\mu_y)]$; Variance $V[X+Y] = V[X]+V[Y]+Cov[X,Y]$

$V[X] = E[X-E(X)]^2 = E[X^2]-(E[X])^2 = E[(X-a)^2] - [E(X)-a]^2$ any a

$V[X+a] = V[X]$; $V[CX] = C^2V[X]$; if independ $V[X+Y] = V[X]+V[Y]$

Approx: if $\mu = E(X)$, $\sigma^2 = V[X]$, $Y=H[X]$: $E[Y] \approx H(\mu) + \frac{1}{2}H''(\mu)\sigma^2$

and $V[Y] \approx [H'(\mu)]^2\sigma^2$

☆ Correlation coeff $\rho_{XY} = \frac{E\{[X-E(X)][Y-E(Y)]\}}{\sqrt{V(X)V(Y)}} = \frac{E(XY)-E(X)E(Y)}{\sqrt{V(X)V(Y)}}$

$(-1 \leq \rho_{XY} \leq 1)$ $\rho = 0$: uncorrelated

If X & Y independ, $\rho_{XY}=0$ (converse only necessarily true for Bernoulli)

If $Y = aX+b$, $\rho = \pm 1$ (converse true)

$E(X|y) = \begin{cases} \sum x_i\, p(x_i|y) \\ \int_{-\infty}^{\infty} x g(x|y)dx \end{cases}$

$E[E(X|Y)] = E[X]$ and $E[E(Y|X)] = E(Y)$

if X & Y independ, $E[X|Y] = E[X]$; regression curve of Y on $X \Rightarrow E(Y|x)$ vs. x.

if regression of Y on X is linear, ρ = correlation coeff and $E(Y|x) = \mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(x-\mu_X)$

if regression of X on Y linear, $E[X|y] = \mu_X + \rho\frac{\sigma_X}{\sigma_Y}(y-\mu_Y)$

Ch8 Special Discrete

(1) Binomial $X \sim Bin(n,p)$ = # successes in n independ trials w/ prob. p; $q = 1-p$,

$P[X=x] = p_X(k) = \binom{n}{k}p^k q^{n-k}$, $E[X] = np$, $V[X] = npq$; $p_X \approx Pois(\alpha = np)$ k passed ? used

(2) Bernoulli: Binomial of $n=1$, $V[X] = pq$, $E[X] = p$ $X = \{0,1\}$

(3) Geometric: $X \sim Geom(p)$ = # trials until success. $p_X(k) = pq^{k-1}$, $E[X] = 1/p$, $V[X] = \frac{1-p}{p^2}$; $E[X^2] = \frac{2}{p^2} - \frac{1}{p}$

MEMORYLESS!

(4) Pascal ("Negative Binomial") $Y \sim Pas(r,p)$ = # trials until r^{th} success.

$p_Y(k) = P[Y=k] = P\left[\begin{matrix}\text{exactly k trials}\\ \text{for } r^{th} \text{ success}\end{matrix}\right] = \binom{k-1}{r-1} p^r q^{k-r}$ $k=r, r+1 \ldots$

$E[Y] = r/p$, $V[Y] = r\left(\frac{1-p}{p^2}\right)$; $P[Y \le n] = P[X^{Binom} \ge r]$

(5) Hypergeometric: simple random sample w/out replace. of n items, N population,

p = proportion "successful" in N: $p_X(k) = \frac{\binom{Np}{k}\binom{N-Np}{n-k}}{\binom{N}{n}}$ $k = 0, 1 \ldots n$

$E[X] = np$; $V[X] = npq\left(\frac{N-n}{N-1}\right)$ ← population correction factor

if $N \gtrsim 20n$, $p_X(k) \approx \binom{n}{k} p^k q^{n-k}$

(6) Multinomial Distribution (multiple outcomes), generalization of Binomial,

pmf $p(n_1, n_2 \ldots n_k) = \frac{n!}{n_1! n_2! \cdots n_k!} p_1^{n_1} p_2^{n_2} \cdots p_k^{n_k}$; $\sum n_i = n$

(7) Poisson Distribution $X \sim Pois(\alpha)$ $p_X(k) = e^{-\alpha} \frac{\alpha^k}{k!}$ $(k = 0, 1, 2 \ldots)$

$E[X] = \alpha$, $V[X] = \alpha$

(Ch 9) Contin. r.v.

(1) Uniform $X \sim Unif(a,b)$; pmf $f(x) = \{\frac{1}{b-a}$ if $a<x<b$; 0 else$\}$

$E[X] = \frac{1}{2}(a+b)$, $V[X] = (b-a)^2/12$

(2) Normal distribution $X \sim Norm(\mu, \sigma^2)$; $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left[\frac{(x-\mu)}{\sigma}\right]^2}$

$E[X] = \mu$, $V[X] = \sigma^2$, $P[a \le X \le b] = P\left[\frac{a-\mu}{\sigma} \le Y \le \frac{b-\mu}{\sigma}\right] = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$

(3) Standard Normal distribution $Z \sim N(\mu=0, \sigma=1)$ has pdf $\varphi = \frac{1}{\sqrt{2\pi}} e^{-1/2 z^2}$

cdf $\Phi(Z) = P[Z \le z] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-1/2 z^2} dt$ {see table}

Thm If $X \sim N(\mu, \sigma^2)$ then $Y = aX + b \sim N(a\mu + b, a^2\sigma^2)$; corr if $X \sim N(\mu, \sigma^2)$, $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$

Time required until 1st occurrence of A if A~Pois(); No memory.

(4) Exponential $X \sim ex(\alpha)$ if pdf $f(x) = \{\alpha e^{-\alpha x}$ if $x > 0$, 0 else$\}$ and $\alpha > 0$.

cdf $F(x) = \{1 - e^{-\alpha x}$ if $x \ge 0$, 0 else$\}$; $E[X] = \frac{1}{\alpha}$; $V[X] = \frac{1}{\alpha^2}$

Ch 12 $f_A = n_A/n$ where n_A = # times of actual success, n = # of trials, $P[|f_A - p| < \epsilon] \ge 1 - \frac{p(1-p)}{n\epsilon^2}$ (swappable) Law of large #s.

Normal Approx. of Binom: $X \sim Bin(n,p)$; $Y = \frac{X - np}{[np(1-p)]^{1/2}} \sim N(0,1)$ so $\lim_{n\to\infty} P[Y \le y] = \Phi(y)$ p close to 1/2.

Corrections for continuity: $P(X=k) \approx P(k - \frac{1}{2} \le X \le k + \frac{1}{2})$; $P[a \le X \le b] \approx P[a - \frac{1}{2} \le X \le \frac{1}{2} + b]$

$X = \sum X_i \Rightarrow$ Central Limit: $Z_n = \frac{X - \sum^n \mu_i}{\sqrt{\sum^n \sigma_i^2}} \sim N(0,1)$; if each X_i same prob, $S = \sum X_i$ $T_n = \frac{S - n\mu}{\sqrt{n}\sigma}$

$P[T_n \le t] = \Phi(t)$