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

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

If $A_1 \dots A_n$ partition Ω, $P(B) = \sum_{i=1}^{n} P(A_i)P(B|A_i) = E[P(B|A)]$

$P(A|B) = \dfrac{P(B|A)P(A)}{P(B)}$ Independence: $P(A \cap B) = P(A)P(B)$, $P(A|B) = P(A)$ Conditional Indep: $P(A \cap B|C) = P(A|C)P(B|C)$, $P(A|B \cap C) = P(A|C)$

n objects: $n!$ perms, $n!/(n-k)!$ k-perms, $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ combos, $\frac{n!}{n_1!n_2!\dots n_r!}$ sized partitions

Discrete $var(X) = E[(X - E[X])^2] = E[X^2] - (E[X])^2 = \sigma_X^2$

$E[g(x)] = \sum_x g(x)p_X(x)$ or $\int g(x)f_X(x)\,dx$ $Y = aX+b \Rightarrow E[Y] = aE[X]+b$, $var(Y) = a^2 var(X)$

$p_X(x) = \sum_y p_{X,Y}(x,y)$ $p_{X,Y}(x,y) = p_Y(y)p_{X|Y}(x|y)$ $p_X(x) = \sum_y p_Y(y)p_{X|Y}(x|y)$

$E[X] = E[E[X|Y]] = \sum_y p_Y(y)E[X|Y=y]$

Independence: $p_{X,Y}(x,y) = p_X(x)p_Y(y)\ \forall x,y$; $E[XY] = E[X]E[Y]$; $var(X+Y) = var(X)+var(Y)$

Discrete uniform $[a,b]$: $p_X(k) = 1/(b-a+1)$, $E[X] = \frac{a+b}{2}$, $var(X) = \frac{(b-a)(b-a+2)}{12}$

Bernoulli: $P(1) = p$, $P(0) = 1-p$, $E[X] = p$, $var(X) = p(1-p)$

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

Geometric: $p_X(k) = (1-p)^{k-1}p$, $E[X] = \frac{1}{p}$, $var(X) = \frac{1-p}{p^2}$

Poisson: $p_X(k) = e^{-\lambda}\frac{\lambda^k}{k!}$, $E[X] = \lambda$, $var(X) = \lambda$. approximates binomial w/ $n >> p$, $\lambda = np$

Continuous $E[X] = \int x f_X(x)\,dx$, $var(X) = \int (x - E[X])^2 f_X(x)\,dx = E[X^2] - (E[X])^2$ $Y = aX+b \Rightarrow$ same as ↑

CDF: $F_X(x) = P(X \le x)$ $F_X(k) = \sum_{i=-\infty}^{k} p_X(i)$ or $F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt$

$p_X(k) = F_X(k) - F_X(k-1)$, $f_X(x) = \frac{d}{dx}F_X(x)$ $Y = aX+b$ is normal if X is normal

If $Y \sim N(0,1)$, $X \sim N(\mu, \sigma^2)$, $Y = \frac{X-\mu}{\sigma}$. $F_X(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$

If $A \in$ real line, $P(X \in A) > 0$, $f_{X|A}(x) = \frac{f_X(x)}{P(X \in A)}$ if $x \in A$, 0 else

$E[X|A] = \int x f_{X|A}(x)\,dx$ $f_X(x) = \sum_i P(A_i)f_{X|A_i}(x)$ $E[X] = \sum_i P(A_i)E[X|A_i]$ if A_i's partition Ω

$f_{X,Y}(x,y) = f_Y(y)f_{X|Y}(x|y)$ $f_X(x) = \int f_Y(y)f_{X|Y}(x|y)\,dy$ $E[X] = \int f_Y(y)E[X|Y=y]\,dy$

$f_{X|Y}(x|y) = \dfrac{f_X(x)f_{Y|X}(y|x)}{f_Y(y)}$ Independence: $f_{X,Y}(x,y) = f_X(x)f_Y(y)\ \forall x,y$

$E[XY] = E[X]E[Y]$, $var(X+Y) = var(X)+var(Y)$

derived distributions: get $F_Y(y) = P(g(X) \le y)$, $f_Y(y) = \frac{d}{dy}F_Y(y)$

More Transforms: $M_X(s) = E[e^{sX}] \Rightarrow E[X^n] = \frac{d^n}{ds^n}M_X(s)\big|_{s=0}$

$Y = aX+b \Rightarrow M_Y(s) = e^{sb}M_X(as)$ X, Y indep $\Rightarrow M_{X+Y}(s) = M_X(s)M_Y(s)$

$p_{X+Y}(x+y) = p_X(x) * p_Y(y)$, $f_{X+Y}(x+y) = f_X(x) * f_Y(y)$ Iterated E: $E[E[X|Y]] = E[X]$

conditional var: $var(X) = E[var(X|Y)] + var(E[X|Y])$

Sums of Random #s of RVs: $Y = \sum X_i$, X_i's have same (μ, σ^2): $E[Y] = \mu E[N]$

$var(Y) = \sigma^2 E[N] + \mu^2 var(N)$ $M_Y(s) = M_N(s)\big|_{e^s = M_X(s)}$ $\rho = \dfrac{cov(X,Y)}{\sqrt{var(X)var(Y)}}$

$cov(X,Y) = E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y]$ independence $\Rightarrow$ 0 cov

Least Squares: $E[(X-c)^2]$ minimized by $c = E[X]$, $E[(X-c)^2|Y=y]$ by $c = E[X|Y=y]$

$E[(X-g(Y))^2]$ by $g(Y) = E[X|Y]$ Least squares linear estimator of X based on Y:

$E[X] + \dfrac{cov(X,Y)}{\sigma_Y^2}(Y - E[Y])$ w/ MSE $(1-\rho^2)var(X)$

Cont. uniform $[a,b]$: $f_X(x) = \frac{1}{b-a}$

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

Exponential: $f_X(x) = \lambda e^{-\lambda x}$ for $x \ge 0$, $F_X(x) = 1 - e^{-\lambda x}$ for $x \ge 0$, $E[X] = \frac{1}{\lambda}$, $var(X) = 1/\lambda^2$

Normal: $\frac{1}{\sigma\sqrt{2\pi}}e^{-(x-\mu)^2/2\sigma^2} = f_X(x)$, $E[X] = \mu$, $var(X) = \sigma^2$