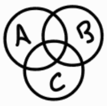
$P(A) = \dfrac{\#\text{ A occur}}{\text{total occur}}$

As $n \to \infty$, accuracy ↑, precision ↑

S = sample space $P(S)=1$

$1 = P(A) + P(A^c)$



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

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

Disjoint: $P(A \cap B) = 0$

Indep: $P(A \cap B) = P(A) \cdot P(B)$

odds $= \dfrac{p}{1-p}$, $\dfrac{0.75}{1-0.75} = 3:1$ odds (wins, loss) can compare odds

$p = \dfrac{0}{0+1}$

Bayes vs Freq — If one-off like cancer, Bayes
prior — norm

Prob dist: set of all outcomes & their probs.

mode: most frequent

mean: $E(X) = \sum_x x \cdot P(X=x)$

entropy: $H(X) = -\sum_x P(X=x) \log(P(X=x))$

var: $Var(X) = E(X^2) - E(X)^2$

$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ $\quad sd(X) = \sqrt{Var(X)}$

$E(aX) = aE(X)$

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

$E(aX+bY) = aE(X) + bE(Y)$

$Var(aX) = a^2 Var(X)$

$Var(X+Y) = Var(X) + Var(Y)$

$Var(aX+bY) = a^2 Var(X) + b^2 Var(Y)$

If X, Y dep, then

$Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$

$P(A|B) = \dfrac{P(A \cap B)}{P(B)}$

If condition causes table to update:

0.1	0.3	0.3	0.3	0.6
X	X	$\frac{0.3}{0.6}$	$\frac{0.3}{0.6}$	
x 1	2	3	4	

$x \geq 2$

Bern: $X \sim Bern(p)$

$P(X=x|p) = p^x(1-p)^{1-x}$

$\mu = p \quad var = p(1-p)$

Binom: $X \sim Binom(n,p)$

$P(X=x|n,p) = \binom{n}{x} p^x (1-p)^{n-x}$ $\quad x = 0,1,\ldots,n$

$\mu = np \quad var = np(1-p)$

Geom: $X \sim Geometric(p)$

$P(X=x|p) = p(1-p)^x \quad x = 0,1,\ldots$

$\mu = \dfrac{1-p}{p} \quad Var = \dfrac{1-p}{p^2}$

Neg binom:

$X \sim Neg\ binom(K, p)$ — X: losses, K: wins

$P(X=x) = \binom{K-1+x}{x} p^K (1-p)^x \quad x = 0,1,\ldots$

$\mu = \dfrac{K(1-p)}{p} \quad var = \dfrac{K(1-p)}{p^2}$

Pois: $X \sim Pois(\lambda)$

fixed time

$P(X=x) = \dfrac{\lambda^x e^{-\lambda}}{x!}$

$\mu = Var = \lambda$

marginal: dist of 1 RV when many

joint: table with intersection probs

sum of joint = 1

combos in joint can't be random b/c they need to obey marginal

Joint → marginal — row sums or col sums

Marginal ~~→~~ Joint unless indep

Measures of dependence:

$Cov(X,Y)$: Joint var & direction

$= E(XY) - E(X)E(Y)$

$\sum x \cdot y \cdot P(X=x \cap Y=y)$

"+"ve : x↑, y↑

"−"ve : x↑, y↓

0 : x↑, y?

Indep → cov = 0

cov = 0 ~~→~~ Indep

Pearson's corr standardizes to σ_x & σ_Y

$\rho_{XY} = \dfrac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$

$-1 \leq \rho_{XY} \leq 1$ — perfect "−"ve; perfect "+"ve

Kendall's τ_K (non-linear monotonic dep.)

measures (x_i, y_i) to (x_j, y_j) instead of (μ_x, μ_y) neither capture parabolas

$-1 \leq \tau_K \leq 1$,

concord: $x_i < x_j$ & $y_i < y_j$ or $x_i > x_j$ & $y_i > y_j$

discord: $x_i > x_j$ & $y_i < y_j$ or $x_i < x_j$ & $y_i > y_j$

$\tau_K = \dfrac{\#\text{concord} - \#\text{discord}}{\binom{n}{2}}$

y \ x	1	2
1	0.1	~~0.3~~
2	0.2	~~0.4~~
	0.3	

→

y \ x	1
1	1/3
2	2/3

$P(Y=y|X=1)$

$E(Y|X=1) = 1 \cdot \frac{1}{3} + 2 \cdot \frac{2}{3} = \frac{5}{3}$

$E(Y) = \sum_x E(Y|X=x) \cdot P(X=x)$

If $H(X) > 0$ but $H(Y) = 0$, X & Y indep

$P(X=x \cap Y=y | Z=z) = P(X=x|Z=z) \cdot P(Y=y|Z=z)$