Statistics Scribe Notes:

4/2/14

Bayes Rule:

There are 2 events: A & B

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

$P(A|B)$ is the posterior probability

$P(A)$ is the prior probability

$P(B|A)$ is the likelihood

$P(B)$ is the marginal

Example:

A: "You like Shrek"

B: "You like Finding Nemo"

$P(A) = 0.6$ Loosely: how many people as a fraction like Shrek?

$P(B|A) = 0.9$

$P(B|\sim A) = 0.3$

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

$= (0.6 \times 0.9)/P(B)$

We know that:

$P(B) = P(A) \times P(B|A) + P(\sim A) \times P(B|\sim A)$

$= 0.6 \times 0.9 + 0.4 \times 0.3$

$= 0.66$

That means $P(A|B) = (0.6 \times 0.9)/0.66$ approximately equals 0.82.

This means that around 82% of people who like Shrek, also like Finding Nemo.

<u>Another example</u>:

G: guilty

D: DNA evidence matched DNA @ scene

Hypothetical trial in New York City:

$P(G) = 1/10{,}000{,}000$ (Probability that any random person committed a crime in NYC. The population is in the denominator.

Expert: "Probability of a False Match is 1 in 1,000,000."

$P(D|{\sim}G) = 1/1{,}000{,}000$ (If you are innocent, there is a 1 in 1,000,000 chance the test says you are guilty)

$P(D|G) = 1$ (If you are guilty, the test is always right)

$P(G|D) = (P(G) \times P(D|G))/P(D)$

$P(D)$ = weighted average

$= P(G) \times P(D|G) + P({\sim}G) \times P(D|{\sim}G)$

$= 1/10{,}000 \times 1 + 9{,}999{,}999/1{,}000{,}000 \times 1/1{,}000{,}000$

$P(D) = 11/10{,}000{,}000$ (11 people would test guilty who were given the test. 1 would be guilty. 10 would be innocent.)

$P(G|D) = (1/10{,}000{,}000 \times 1)/(1/10{,}000{,}000) = 1/11$ (1 in 11 people would be guilty and test positive in the DNA test)

Confusing $P(G|D)$ with $P(D|G)$ is known as the base rate fallacy or the prosecutor's fallacy.

<u>Odds</u>: Just a different way of expressing probabilities.

A: event that "Robert Apprentice" wins Kentucky Derby

$P(A) = 0.2$

$\text{Odds}(A) = P(A)/(1-P(A))$ = Probability of A/Probability of Not A

$= 0.2/(1-0.2) = 0.2/0.8 = ¼$

$\text{Odds Against}(A) = 1/\text{Odds}(A)$

$= (1 – P(A))/P(A) = 4/1$

<u>Random Variables</u>:

X is a random variable

A probability distribution is a description of the possible outcomes for X, together with their probabilities.

X : coin flip

Outcomes	**Probabilities**
Heads	0.5
Tails	0.5

Distribution as: $P(X = x_k) = w_k$

Probability (Mass) Table

Value (x_k)	**Probability (w_k)**
1	1/9
2	2/9
3	1/9
4	2/9
5	1/9
6	2/9

All probabilities must add up to 1.

There are "discrete" distributions.

<u>Expected value</u>: the average outcome of a R.V.

$$E(X) = \sum_{k=1}^{D} w_k x_k$$

X = price of AAPL

Value	**Probability**
400	0.2
500	0.6
600	0.2

E(X) = 400 x 0.2 + 500 x 0.6 + 600 x 0.2

This is not the simple average, but the weighted average.

E(f(X)) =/= f(E(X)) in general → known as Jensen's inequality

X = AAPL stock price

Y = f(X) = 0.75x

Event x_k	**Probability w_k**	**Function (Event) $f(x_k)$**
400	0.2	0.75 x 400 = 300
500	0.6	0.75 x 500 = 375
600	0.2	0.75 x 600 = 450

$$E(f(x)) = \sum_{k=1}^{D} w_k f(x_k)$$

E(f(x)) = 0.2 x 300 + 0.6 x 375 + 0.2 x 450 = 435

Fact: f(X) = aX + B

Then E(f(X) = a(E(X) + b = f(E(x))

__**Example**__:

Probability	**Price of Copper (x)**	**Copper Demanded (y)**
0.25	2	1000 x 2^{-2}
0.50	3	1000 x 3^{-2}
0.25	4	1000 x 4^{-2}

Y = K x^{β} = 1000x^{-2}

Y = Copper demanded

X = Price of copper

What is E(Y)?

E(Y) = E(f(x) where f(x) = 1000x^{-2}

Using the sigma function above: E(Y) = 0.25 x 1000 x 2^{-2} + 0.50 x 1000 x 3^{-2} + 0.25 x 1000 x 4^{-2}

<u>Variance</u>:

X = Random variable with distribution P

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

$\mu = E(X)$

$var(X) = E((x - u)^2) = E(f(X))$ where $f(X) = (x - u)^2$

$var(X) = E(x^2) - [E(x)]^2$

$= E(x^2) - \mu^2$

$Sd(X) = sqrt[var(x)]$