

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.

**Another example:**

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.)

$P(G|D) = (1/10,000,000 \times 1)/(11/10,000,000) = 1/11$  (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.

**Odds:** Just a different way of expressing probabilities.

A: event that "Robert Apprentice" wins Kentucky Derby

$P(A) = 0.2$

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

$$= 0.2/(1-0.2) = 0.2/0.8 = 1/4$$

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

$$= (1 - P(A))/P(A) = 4/1$$

**Random Variables:**

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:  $P(X = x_k) = w_k$

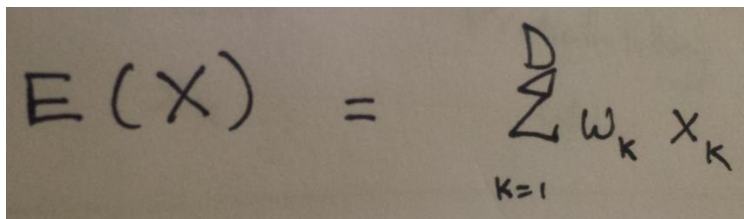
Probability (Mass) Table

| 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.

**Expected value:** the average outcome of a R.V.



A handwritten formula on a piece of paper showing the expected value of a discrete random variable X. The formula is  $E(X) = \sum_{k=1}^D w_k x_k$ , where the summation is from k=1 to D, w\_k represents the probability, and x\_k represents the value.

X = price of AAPL

| Value | Probability |
|-------|-------------|
| 400   | 0.2         |
| 500   | 0.6         |
| 600   | 0.2         |

$$E(X) = 400 \times 0.2 + 500 \times 0.6 + 600 \times 0.2$$

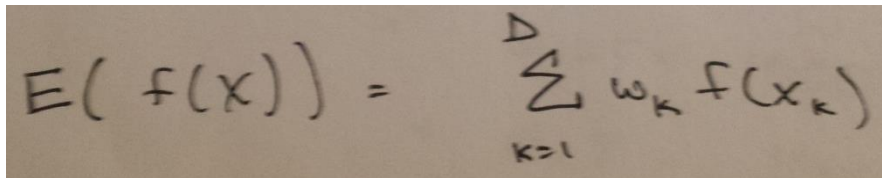
This is not the simple average, but the weighted average.

$E(f(X)) \neq f(E(X))$  in general  $\rightarrow$  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               | $0.75 \times 400 = 300$   |
| 500         | 0.6               | $0.75 \times 500 = 375$   |
| 600         | 0.2               | $0.75 \times 600 = 450$   |



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

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

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                   | $1000 \times 2^{-2}$ |
| 0.50        | 3                   | $1000 \times 3^{-2}$ |
| 0.25        | 4                   | $1000 \times 4^{-2}$ |

$$Y = K x^{\beta} = 1000x^{-2}$$

$Y$  = Copper demanded

$X$  = Price of copper

What is  $E(Y)$ ?

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

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

**Variance:**

$X$  = Random variable with distribution  $P$

$$\text{var}(X) = E([X - E(X)]^2)$$

$$\mu = E(X)$$

$$\text{var}(X) = E((x - u)^2) = E(f(X)) \text{ where } f(X) = (x - u)^2$$

$$\text{var}(X) = E(x^2) - [E(x)]^2$$

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

$$\text{Sd}(X) = \sqrt{\text{var}(x)}$$