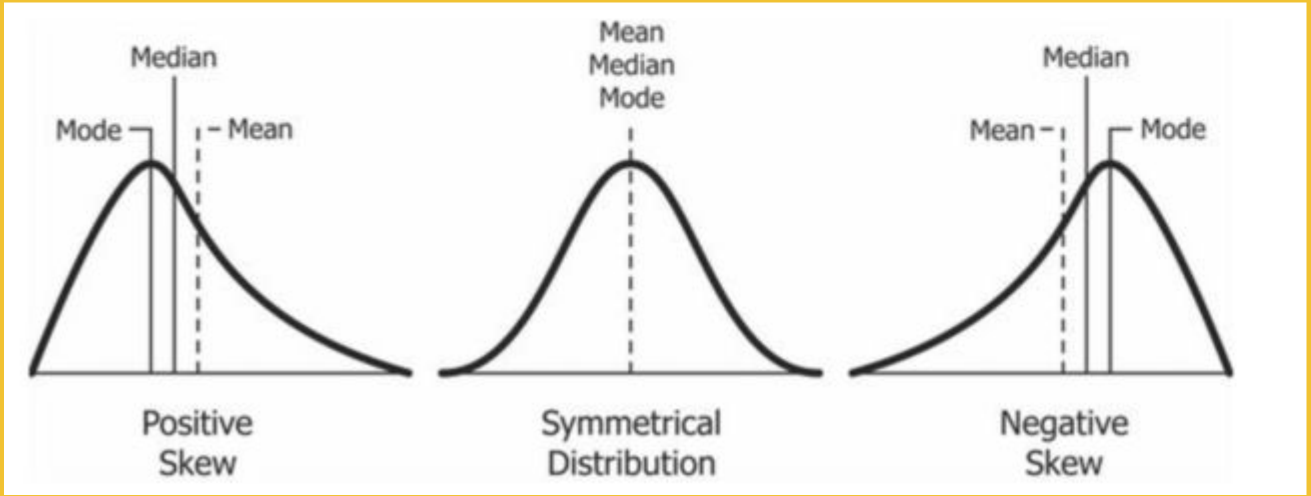



# Econ 2250: Stats for Econ

[Source for pic stats above.]

## Announcements

- Homework 3 is due Sunday

## What we will do today?

- Random variables
- Unconditional probability rules
- Venn Diagrams
- Crosstab in Pandas

# Probability notation

- Capital letters, such as X, Y, and Z, are used to denote random variables.
- Lowercase letters, such as x, y, z and a, b, c are used to denote particular values that the random variable can take on.
- Thus, the expression $P(X = x)$ symbolizes the probability that the random variable X takes on the particular value x. Often, this is written simply as $P(x)$.
- Likewise, $P(X \leq x)$ = probability that the random variable X is less than or equal to the specific value x;
- $P(a \leq X \leq b)$ = probability that X lies between values a and b.

## Properties of a Random Experiment

- An experiment or observation for which there is a known set of possible outcomes.
- But for which we cannot predict the future outcomes.
- An random experiment is repeated under identical conditions, the outcomes or results may fluctuate or vary randomly.

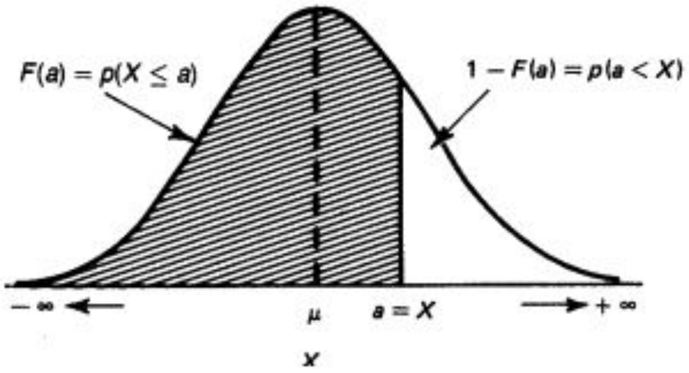
# Properties of a Random Variable

- Result of a random experiment that takes on real values.
- Usually refer to an a set of RV X, and an observation as x.
- Multiplicative: c*X is an RV, where c is a constant.
- Additive: If $X_1$ and $X_2$ are two R.V.s, then $X_1 + X_2$ and $X_1$ $X_2$ are also R.Vs.
- Can be continuous or discrete.

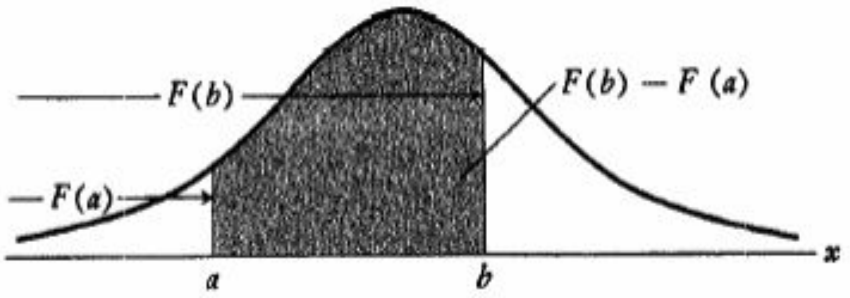
**Probability Distribution Function (PDF): specification of the probability associated with each value of a random variable.**

For continuous r.v.s:

$$F(a) = p(X \le a) = \int_{-\infty}^{a} f(x)\,dx = \text{Area up to } X = a$$




**$p(a \le X \le b) = F(b) - F(a)$**



# Probability Mass Function (PMF)

A probability distribution involving only discrete values of X. Aggregates different possible values of X, and the different possible values of P(x).

| x | P(x) |
|---|---|
| 2 | 1/16 |
| 3 | 2/16 |
| 4 | 3/16 |
| 5 | 4/16 |
| 6 | 3/16 |
| 7 | 2/16 |
| 8 | 1/16 |


Properties:

$0 \leq P(X = x) \leq 1$

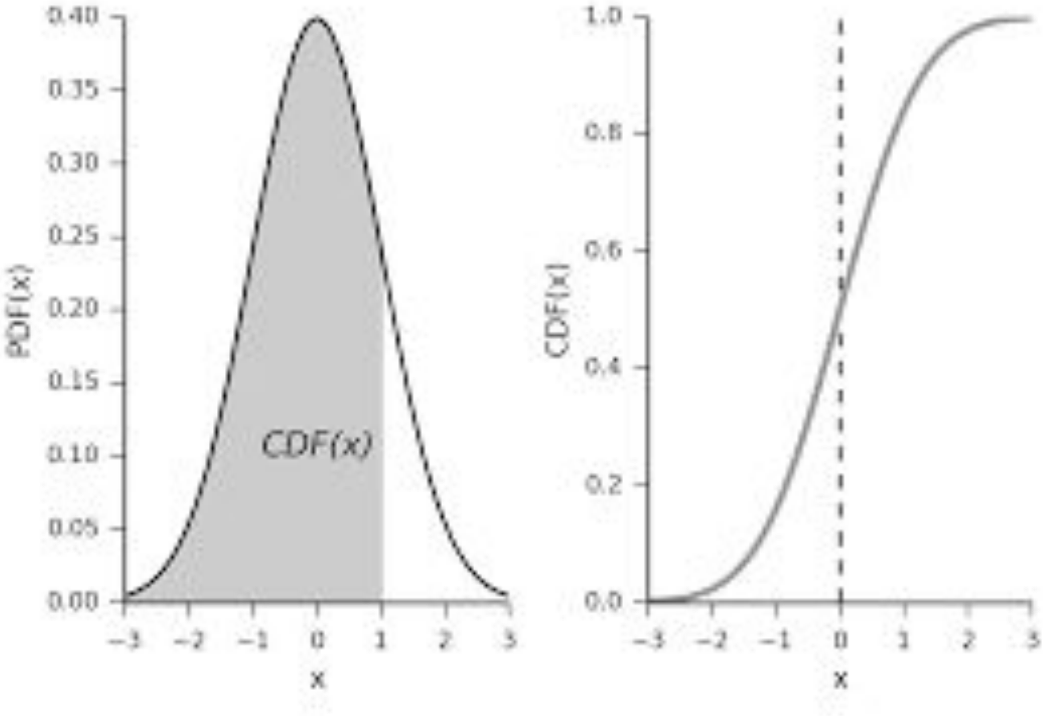
$\Sigma\, P(X = x) = 1.$

# Cumulative Density Function (CDF)

The probability that a random variable X takes on a value less than or equal to some particular value a is often written as

$$F(a) = p(X \le a) = \sum_{X \le a} p(x)$$

(for discrete variables, integral for continuous)



# Basic Rules of Probability

1. For any event P(E) [0,1]
2. If an event cannot occur P(E) = 0
3. If an event is certain to occur P(E) =1
4. The sum of the probability of all outcomes must equal 1.

**Summary of probabilities**

| Event | Probability |
| --- | --- |
| A | $P(A) \in [0,1]$ |
| not A | $P(A^{\complement}) = 1 - P(A)$ |
| A or B | $P(A \cup B) = P(A) + P(B) - P(A \cap B)$<br>$P(A \cup B) = P(A) + P(B) \qquad \text{if A and B are mutually exclusive}$ |
| A and B | $P(A \cap B) = P(A\|B)P(B) = P(B\|A)P(A)$<br>$P(A \cap B) = P(A)P(B) \qquad \text{if A and B are independent}$ |
| A given B | $P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B\|A)P(A)}{P(B)}$ |

# Probability Jargon

**Marginal Probability**: P(A)

**Joint Probability**: P(A and B) = P(A,B)

**Conditional Probability**: P(A given B) = P(A|B)

P(A|B) = P(A,B)/P(B)

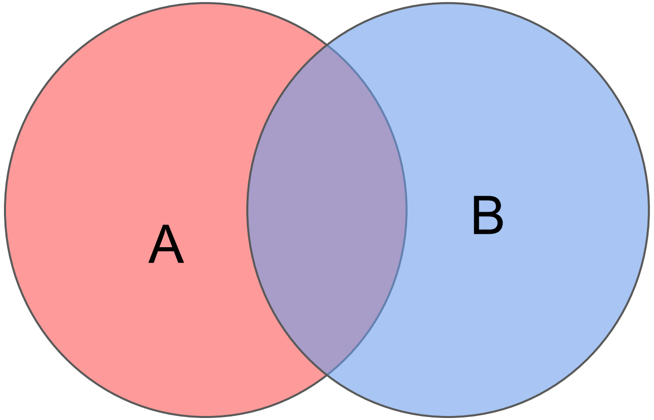
**NOTICE: P(A|B)  NOT EQUAL P(B|A)**

Likelihood of event

$$P(\text{event}) = \frac{\#\text{ of outcomes of event}}{\#\text{ of outcomes in }\Omega}$$

# Events

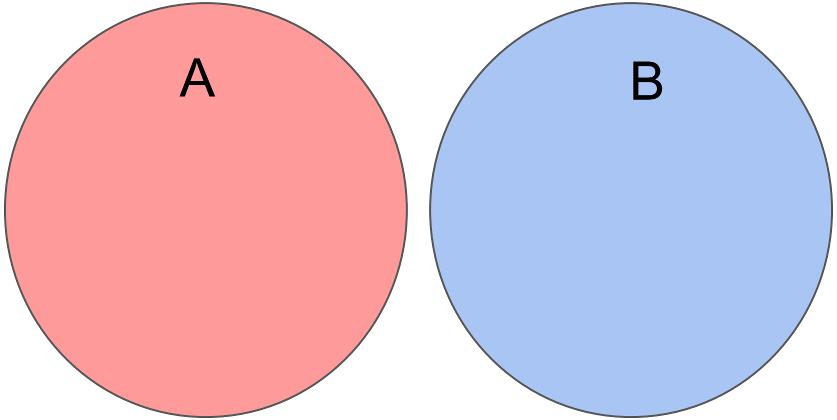
## Non-mutually exclusive




P(A U B)

P(A) + P(B) - P(A|B)

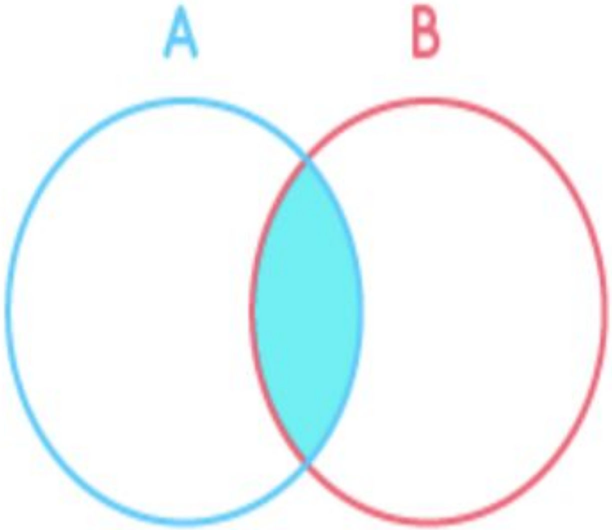
## Mutually exclusive




P(A U B)

P(A) + P(B)

Independent




$P(A \cap B)$

P(A) * P(B)

# Bayes Rule

- P(A|B) = P(B|A) * P(A) / P(B)
- NOTE: we often do not have access to P(B) and have to calculate by looking at all possible cases:
- P(B) = P(B|A) * P(A) + P(not B|not A) * P(not A)
    a. P(not A) = 1 – P(A)
    b. P(not B|not A) IS UNKNOWN, needs to be given

# Here are some examples of bayes rule

https://www.mathsisfun.com/data/bayes-theorem.html

Make sure to check out the test questions at the bottom. You should be able to identify

P(A|B) (what you're looking for),

P(B|A) the prior,

P(A) the marginal of the conditional that you are looking for, and

P(B) marginal of the condition (or how to find it)

If no P(B), define P(A)P(B|A) + P(not A)P(not A|not B)

# Bayes Rules example 1

Given inputs, what is the probability of dangerous Fire when there is Smoke?

- dangerous fires are rare (1%)
- but smoke is fairly common (10%) due to barbecues,
- and 90% of dangerous fires make smoke

$P(\text{Fire} \mid \text{Smoke})?$

$P(\text{Fire}) = 1\%$

$P(\text{Smoke}) = 10\%$

$P(\text{Smoke} \mid \text{Fire}) = 90\%$

$$P(\text{Fire} \mid \text{Smoke}) = P(\text{Fire}) * P(\text{Smoke} \mid \text{Fire}) / P(\text{Smoke}) =$$

$$(1\% * 90\%) / 10\% = 9\%$$

# Bayes example 2

At a School, 60% of the boys play football and 36% of the boys play ice hockey.

Given that 40% of those that play football also play ice hockey, what percent of those that play ice hockey also play football?

P(A|B)?

P(A) = 60% = 0.6

P(B) = 36% = 0.36

P(B|A) = 40% = 0.4

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)} = \frac{0.6 \times 0.4}{0.36} = \frac{0.24}{0.36} = 66\tfrac{2}{3}\%$$

# Bayes Example 3

Dr. Foster remembers to take his umbrella with him 80% of the days.

It rains on 30% of the days when he remembers to take his umbrella, and it rains on 60% of the days when he forgets to take his umbrella.

What is the probability that he remembers his umbrella when it rains?

# P(Remembers | Rains)

P(Remembers) = 80%

P(Rains | Remembers) = 30%

**R(Rains) = 30% × 80% + 60% × 20% = 0.24 + 0.12 = 0.36**

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)} = \frac{0.8 \times 0.3}{0.36} = \frac{0.24}{0.36} = \frac{2}{3}$$

# Bayes Example 4

In a factory, machine X produces 60% of the daily output and machine Y produces 40% of the daily output.

2% of machine X's output is defective, and 1.5% of machine Y's output is defective.

One day, an item was inspected at random and found to be defective. What is the probability that it was produced by machine X?

P(X|defective)

P(X) = 60% = 0.6

P(defective) = 2% × 60% + 1.5% × 40% = 0.012 + 0.006 = 0.018

P(defective|X) = 2% = 0.02

$$\frac{P(A)P(B|A)}{P(B)} = \frac{0.6 \times 0.02}{0.018} = \frac{0.012}{0.018} = \frac{2}{3}$$

# End of class form

https://forms.gle/QURA6pDK4MZLrqpp9