## Unit 4:

### Conditional Probability:

$$P(\text{video games} \mid \text{girl}) = \frac{12}{60} = .20$$

**event**
The event is always contained within the conditional sample space.

**condition**
The condition is always just a portion of the sample space (the *conditional sample space*).

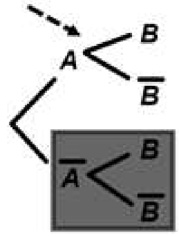
The event goes in the numerator of the fraction.

$$P(\text{video games} \mid \text{girl}) = \frac{12}{60} = .20$$

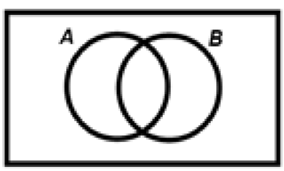
The condition goes in the denominator of the fraction.

The conditional sample space is a portion of the sample space
The event is a portion of the conditional sample space.

| | girls | boys | |
|---|---|---|---|
| read a book | 18 | 4 | 22 |
| video games | 12 | 20 | 32 |
| watch Netflix | 30 | 16 | 46 |
| | 60 | 40 | 100 |

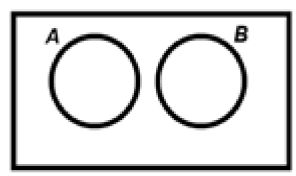



OR is add (but must subtract any overlap):




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

Special Case for OR:
Disjoint (Mutually-Exclusive) Events




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

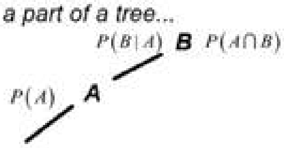
*So the OR formula is simplified...*

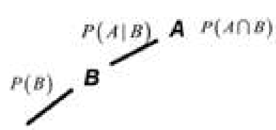
$$P(A \cup B) = P(A) + P(B)$$

AND is multiply (but 2nd probability must be conditional):

$$P(A \cap B) = P(A) \cdot P(B \mid A) \qquad P(A \cap B) = P(B) \cdot P(A \mid B)$$

*Picture a part of a tree...*




Special Case for AND: Independent Events

**Test for independent events:**

Two events are independent if:

$$P(B) = P(B \mid A) = P(B \mid \overline{A})$$

(check any two)

For independent events: $P(B) = P(B \mid A)$

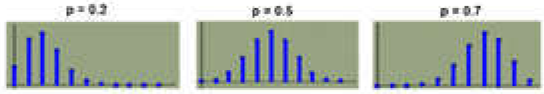
$$P(A \cap B) = P(A) \cdot P(B \mid A) \quad \text{...simplifies to...}$$

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

## 8: Discrete Probability Models

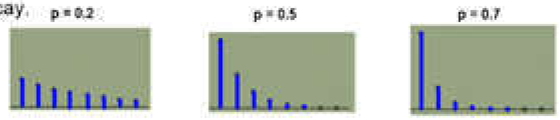
**Binomial** Shape depends upon p.

$\mu = np \quad \sigma = \sqrt{npq}$




**Geometric** Shape is always exponential decay.

$\mu = \frac{1}{p} \quad \sigma = \frac{\sqrt{1-p}}{p}$



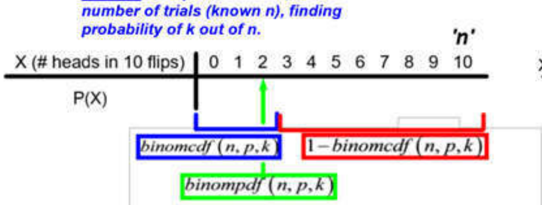

**General Discrete Models**

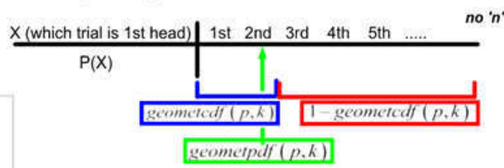
$\mu = \text{'expected value'} = \sum X \cdot P(X)$

$\sigma \text{ (and } \mu\text{) found using L1(data), L2(freqList), 1-Var Stats}$

## 8: Discrete Probability Models

| **Binomial** | **Geometric** |
|---|---|
| - Only 2 outcomes | - Only 2 outcomes |
| - Probabilities must be the same each trial. | - Probabilities must be the same each trial. |
| - Probabilities of trials must be independent. | - Probabilities of trials must be independent. |
| - Must have fixed number of trials, n | - May or may not have fixed number of trials, n |
| ***Best for**: independent trials, fixed number of trials (known n), finding probability of k out of n.* | ***Best for**: independent trials, non-fixed number of trials (unknown n), finding probability of 'when' the 1st success occurs.* |

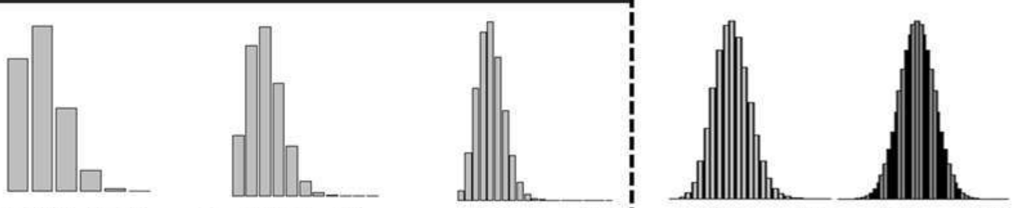



$P(\text{exactly } k \text{ successes out of } n \text{ trials}) = {}_nC_k (p)^k (q)^{n-k}$

$P(\text{success on the } k^{th} \text{ trial}) = (q)^{k-1}(p)$

## 10: Normal Approximation of Binomial Model



$p = 0.2, n = 5$ $(np = 1)$ | $p = 0.2, n = 10$ $(np = 2)$ | $p = 0.2, n = 20$ $(np = 4)$ | $p = 0.2, n = 50$ $(np = 10)$ | $p = 0.2, n = 100$ $(np = 20)$

**Can use Normal approximation for the Binomial distribution.**

*If* $np \geq 10$ *and* $nq \geq 10$

*a Binomial distribution can be approximated with a Normal distribution with:* $\mu = np$

$\sigma = \sqrt{npq}$

## 11: Combining Multiple Distributions

Define an algebraic expression for how the source distributions are used to build the new distribution:

$E = A + B - C - D$

***The means are always determined by <u>the defining algebraic expression</u>:***

$\mu_E = \mu_A + \mu_B - \mu_C - \mu_D$

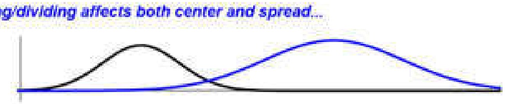
***But because each source of variability increases overall variation, the variances always add:***

$\sigma_E^2 = \sigma_A^2 + \sigma_B^2 + \sigma_C^2 + \sigma_D^2$

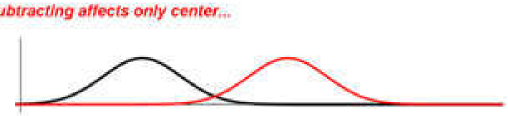
***However, we must know for certain that the variables are all varying independently of one another. (If not independent, we can find mean but not standard deviation).***

## 11: Transforming a Single Distribution

*Multiplying/dividing affects both center and spread...*



*Adding/Subtracting affects only center...*



*If* $Y = aX \pm b \qquad \mu_Y = a\mu_X \pm b \qquad \sigma_Y = a\sigma_X$