# Discrete Probability Distributions

Math 122

# Random Variables

- A variable x whose value is determined by the outcome of an experiment.
- P(x) = probability of a particular value of x
- Mean/Expected Value: $\mu = \sum xP(x)$
  - If the experiment is repeated many times and the values of x are averaged, the average should be near $\mu$.
- Standard Deviation: $\sigma$

# Range Rule of Thumb

- Usual values are between

$$\mu - 2\sigma \text{ and } \mu + 2\sigma$$

# 5% Rule

- If $P(x \leq N) \leq 5\%$, then N is unusually low

- If $P(x \geq N) \leq 5\%$, then N is unusually low

# Special Distributions

- We want a few specific, common distributions so that we know what to do when we encounter them.
- Discrete
  - Binomial (counting successes in trials)
  - Poisson (counting events in an interval)
- Continuous
  - Uniform (simple)
  - Normal (pervasive bell curve)
  - t, F, $\chi^2$

# Binomial Distribution

# Binomial Distribution

- A fixed number of trials is repeated.
- The trials are independent.
- Each trial ends in success or failure.
- The probability of success is the same for each trial.
- The value of x is the number of successes.

# Examples of Binomial Distributions

- The number of female children out of 10 randomly selected children
- The number of green peas in sets of 5 offspring peas.
- The number of correct responses when you guess on a multiple choice test.
- The number of Republicans among sets of 1000 random voters.

# Binomial Distribution Pre-Example

What is the probability of getting exactly 3 T/F questions correct when you guess on 4?

# Binomial Distribution Notation

- n = number of trials
- p = probability of success
- q = probability of failure = 1-p
- P(x) = probability of getting exactly x successes in n trials.

# A formula we won't use

$$P(x) = \frac{n!}{(n-x)!x!} p^x q^{n-x}$$

# Functions we will use

- P(x = N) = binom**p**df(n, p, N)
- P(x ≤ N) = binom**c**df(n, p, N)
- “c” is for “cumulative”

What is the probability that a family with 5 children has exactly 3 boys?

What is the probability that a family with 5 children has 3 or fewer boys?

What is the probability that a family with 5 children has at least 3 boys?

# Lesser/Greater
# What is the probability that x is…

- equal to N? pdf (N)
- less than or equal to N? cdf (N)
- less than N? cdf(N – 1)
- greater than or equal to N? 1 – cdf (N – 1)
- greater than N? 1 – cdf (N)

# Lesser/Greater

- **at most** means **less than or equal**
- **at least** means **greater than or equal**
- **no more than** means **less than or equal**
- **no less than** means **greater than or equal**
- **up to** means **less than or equal**

# Guessing

- A true/false test has 100 questions. Each question has 2 options, of which one is correct. You guess on every question.
  - What is the probability that you get at least half correct?
  - What is the probability you get no more than 30 correct?
  - What is the probability that you get 60 or more correct?
  - What is the probability that you get exactly 50 correct?
  - What is the probability that you get 100 correct?

# Another 5% Rule

- In sampling without replacement, the individuals are not independent, but…

- If the sample size is no more than 5% of the population, then we can treat the individual as independent.

On a college campus of 10,000 students, 2/3 of the students are female. If 10 students are chosen at random, what is the probability that no more than half of them are female?

# Binomial Mean and Standard Deviation

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

# Gregor Mendel

- Gregor Mendel Estimated that the probability that a pea pod with green/yellow genes turns out to be green is 0.75.
- To test his claim, he bred 580 pea pods. Of these, 428 were green.
- If Mendel was correct, would this be unusual (according to the Range Rule of Thumb)?

# Gregor Mendel

- Gregor Mendel Estimated that the probability that a pea pod with green/yellow genes turns out to be green is 0.75.
- To test his claim, he bred 580 pea pods. Of these, 428 were green.
- If Mendel was correct, would this be unusual (according to the 5% Rule)?

# Racial Discrimination

- 79.1% of the population of Hidalgo County, TX, is of hispanic descent.
- Of 870 people selected for jury duty for a case of burglary against Rodrigo Partida, 339 or 39% were hispanic.
- After conviction, Partida was granted a new trial because of the discrepancy of 39% compared to 79%.
- Statistically, would 339 of 870 be an unusually low number in this case?

# Lottery

- In the Illinois Pick 3 Lottery game, players select three digits. There is one winning sequence among the 1000 possibilities.
- If you play this game every day for a year, what is the probability that you win at least once?

# Poisson Distribution

# Poisson Distribution

- The Poisson distribution is a discrete probability distribution that applies to occurrences of some event over a specified interval.
- The interval can be time, distance, area, volume, or some similar unit.
- The random variable x is the number of occurrences of the event in an interval.

# Example Poisson Distributions

- The number of major earthquakes during a year.
- The number of births at a hospital in a year.
- The number of emails received in an hour.
- The number of automobile accidents on a given mile of road.
- The number of bug pieces in a tablespoon of peanut butter.
- The number of dandelions on a square foot of dirt.

# More Formulas We Won't Use

For a Poisson Distribution with mean $\mu$

$$\sigma = \sqrt{\mu}$$

$$P(x) = \frac{\mu^x \cdot e^{-\mu}}{x!}$$

Where $e \approx 2.718281828459045$

# Functions we will use

# Functions we will use

- P(x = N) = poisson**p**df(n, p, N)
- P(x ≤ N) = poisson**c**df(n, p, N)
- “c” is for “cumulative”

# Earthquakes

- According to the USGS, there have been 16,500 earthquakes at or above magnitude 6 since 1900.
- What is the probability that there are 125 or fewer earthquakes at or above magnitude 6 in a given year?

# Earthquakes

- What would be the usual range for the number of earthquakes at or above magnitude 6 during one year (according to the Range Rule of Thumb)?

# Births

- 120 children are born each year at Seward Memorial Hospital.
- SMH has 2 “birthing rooms.”
- What is the probability on any given day that this is adequate?

# Bugs in Peanut Butter

- The USDA allows a maximum of 30 “insect parts” in 100g or 3.53oz of peanut butter (twice that for chocolate).
- Suppose that a jar of peanut butter has the maximum allowable number of bug parts.
- If a sandwich is made with one ounce of this peanut butter, then what is the probability that the peanut butter in the sandwich does not contain any bug parts?

# Bugs in Peanut Butter

- The USDA allows a maximum of 30 “insect parts” in 100g or 3.53oz of peanut butter (twice that for chocolate).
- Suppose that a jar of peanut butter has the maximum allowable number of bug parts.
- If a sandwich is made with one ounce of this peanut butter, then what is the probability it contains at least 5 bug parts?

# Tornadoes

- There were 7236 tornadoes in Texas in a recent span of 54 years (the most of any state in the USA).
- What is the probability that there are 110 or fewer tornadoes in one year in Texas?
- What is the probability that there are more than 150 tornadoes in one year in Texas?