Data

• We’ve talked about two broad types of data this semester
• Do you remember what they are?
Probability

• We learned several rules for calculating probabilities of events
  – E.g. the event ‘Tails’ in the flip of a fair coin
• Special relationships between two events
  – Mutually exclusive
  – Independent
  – Could you provide definitions and examples of each?
Random Variables

• We’ve talked about two types of random variables
  – Can you name them?
• We’ve talked about three different models (i.e. distributions) for random variables
  – Can you describe them (be specific)?
  – Can you find probabilities under each?
  – Do you know which model is appropriate for each type?
Sampling Distribution of p-hat

• Can you describe what the sampling distribution of p-hat is?

• Do you know when to use it?
  – When the problem asks you to find a probability for a sample proportion
  – When the sample size is large enough: \( np \geq 10, \ n(1-p) \geq 10 \)
Confidence Intervals for $p$

• Useful when we want to provide a range of reasonable estimates for the population proportion

• Do you know how to interpret the confidence level?

• Do you know how to use a confidence interval to make a decision, and under what conditions this would be appropriate?
Hypothesis Testing for $p$

- Useful for making a Yes/No type decision about particular value(s) of a population proportion
- Do you know the 5 steps of hypothesis testing?
- Can you interpret and calculate a test statistic?
- Can you interpret and calculate a p-value?
If your p-value is significant at the 5% level, then it will also be significant at the …

A) 1% level  
B) 10% level  
C) Both of the above  
D) None of the above
True or False

The standard error of p-hat estimates, roughly, the average distance of an observed p-hat value from the true population proportion.
I tell you “A and B are mutually exclusive.” Choose all of the true statements below.

A) A and B are independent
B) \( P(A \text{ and } B) = 0 \)
C) \( P(A \text{ and } B) = P(A)P(B) \)
D) \( P(A \text{ or } B) = P(A) + P(B) \)
E) None of the above are true
X has the Bin(n,p) distribution. How would you find $P(\hat{p} > 0.2)$? Choose all that apply...

A) Use sampling distribution of $\hat{p}$ to find $P(\hat{p} > 0.2)$ directly

B) Use normal approximation to binomial to find $P(X > 0.2*n)$ instead

C) Not enough information
Which of the following distributions for X meet the criteria to use the normal approximation? Choose all that apply...

- Binomial(25, 0.2)
- Binomial(25, 0.5)
- Binomial(25, 0.7)
- None of the above
Which graph(s) would be appropriate to display a quantitative variable? Choose all that apply...

Bar chart

Histogram

Boxplot

QQ-plot

Time-series plot
Which graph(s) would be appropriate to display shape of a distribution? Choose all that apply...

Histogram

Boxplot

QQ-plot

Time-series plot
A well constructed survey of monkeys found that 75% of monkeys with prehensile tails prefer bananas to grubs. Express this as a conditional probability.

\[ P(\text{prehensile tail} \mid \text{prefer bananas}) = 0.75 \]

\[ P(\text{prefer bananas} \mid \text{prehensile tail}) = 0.75 \]
To determine if monkey tail status (prehensile or not) is independent of food preference (bananas or grubs), we could compare

$P(\text{prefer bananas} \mid \text{prehensile tail})$ to...

$P(\text{prehensile tail} \mid \text{prefer bananas})$

$P(\text{prehensile tail})$

$P(\text{prefer bananas})$
A 95% confidence interval will be ________ a 99% confidence interval.

Narrower than

Wider than

The same as

Can’t tell
Match the type of random variable to the appropriate distribution(s)

1. Discrete
   a. Normal(μ,σ)

2. Continuous
   b. Uniform(a,b)
   c. Binomial(n,p)
True or False

A z-statistic tells you how many standard errors your observed sample proportion is from the true proportion.