MA145 Final Exam Review

Gene Quinn

Binomial Experiment

A **binomial experiment** consists of:

- A fixed number n of independent trials with two outcomes, success and failure

Binomial Experiment

A **binomial experiment** consists of:

- A fixed number n of independent trials with two outcomes, success and failure
- Each trial has probability p of success and probability $1 - p$ of failure

Binomial Random Variable

A **binomial random variable** X is defined as the number of successes in a binomial experiment with n trials.

Binomial Random Variable

A **binomial random variable** X is defined as the number of successes in a binomial experiment with n trials.

The **mean of a binomial variable** is

$$\mu_X = np$$

Binomial Random Variable

A **binomial random variable** X is defined as the number of successes in a binomial experiment with n trials.

The **mean of a binomial variable** is

$$\mu_X = np$$

The **standard deviation of a binomial variable** is

$$\sigma_X = \sqrt{np(1-p)}$$

Binomial Experiment Examples

Example: A fair coin is tossed 5 times.

Binomial Experiment Examples

Example: A fair coin is tossed 5 times.

This is a binomial experiment because we have a fixed number of trials (5), each with two outcomes.

If we (arbitrarily) identify the outcome $heads$ with "success", the probability of success is constant in each trial.

Binomial Experiment Examples

Example: A fair coin is tossed 5 times.

This is a binomial experiment because we have a fixed number of trials (5), each with two outcomes.

If we (arbitrarily) identify the outcome $heads$ with "success", the probability of success is constant in each trial.

Based on the statement that the coin is fair, our binomial experiment has:

$$n = 5 \quad \text{and} \quad p = \frac{1}{2}$$

Binomial Experiment Examples

Example: We draw $n = 10$ random samples from a normal population and calculate 10 95% confidence intervals for the population mean μ.

Binomial Experiment Examples

Example: We draw $n = 10$ random samples from a normal population and calculate 10 95% confidence intervals for the population mean μ.

This is a binomial experiment because we have a fixed number of trials (10), each with two outcomes:

- The confidence interval contains the population mean μ (success)
- The confidence interval does not contain the population mean μ (failure)

Binomial Experiment Examples

Example: We draw $n = 10$ random samples from a normal population and calculate $10\ 95\%$ confidence intervals for the population mean μ.

This is a binomial experiment because we have a fixed number of trials (10), each with two outcomes:

- The confidence interval contains the population mean μ (success)
- The confidence interval does not contain the population mean μ (failure)

Based on the fact that the probability that a 95% confidence interval contains the population mean is 0.95, the probability of success is $p = 0.95$.

our binomial experiment has:

$$n = 10 \quad \text{and} \quad p = 0.95$$

The Law of Large Numbers

The **Law of Large Numbers** says that, as we add observations to a sample, the difference between the sample mean and the population mean tends to zero:

$$\overline{x} - \mu \longrightarrow 0 \quad \text{as } n \text{ becomes large}$$

The Law of Large Numbers

The **Law of Large Numbers** says that, as we add observations to a sample, the difference between the sample mean and the population mean tends to zero:

$$\overline{x} - \mu \to 0 \quad \text{as } n \text{ becomes large}$$

Picture an experiment where we draw observations one at a time from a population, and after each draw we recalculate the mean including the new observation.

The Law of Large Numbers says that if we keep doing this long enough, the difference between the sample mean $\overline{x}$ and the population mean μ will become very small.

The Law of Large Numbers

The **Law of Large Numbers** says that, as we add observations to a sample, the difference between the sample mean and the population mean tends to zero:

$$\overline{x} - \mu \rightarrow 0 \quad \text{as } n \text{ becomes large}$$

Picture an experiment where we draw observations one at a time from a population, and after each draw we recalculate the mean including the new observation.

The Law of Large Numbers says that if we keep doing this long enough, the difference between the sample mean $\overline{x}$ and the population mean μ will become very small.

The Central Limit Theorem

The **Central Limit Theorem** deals with the sampling distribution of $\overline{x}$, the mean of a random sample.

The Central Limit Theorem

The **Central Limit Theorem** deals with the sampling distribution of $\overline{x}$, the mean of a random sample.

As stated in the text, the theorem assumes that the sample is taken from a population with mean μ and standard deviation σ

The Central Limit Theorem

The **Central Limit Theorem** deals with the sampling distribution of $\overline{x}$, the mean of a random sample.

As stated in the text, the theorem assumes that the sample is taken from a population with mean μ and standard deviation σ

The theorem says that, even if the data are *not* normally distributed, the sampling distribution of $\overline{x}$ becomes approximately normal with

mean of $\overline{x} = \mu_X = \mu$ and standard deviation of $\overline{x} = \sigma_{\overline{x}} = \dfrac{\sigma}{\sqrt{n}}$

The Central Limit Theorem

The importance of the Central Limit Theorem is that it allows us to treat the sample mean $\overline{x}$ as a normally distributed variable, even if the data is not normally distributed.

The Central Limit Theorem

The importance of the Central Limit Theorem is that it allows us to treat the sample mean $\overline{x}$ as a normally distributed variable, even if the data is not normally distributed.

This is why our tests are stated in terms of z and t values, which arise from sampling normal populations, even when the underlying populations are not normally distributed.

Statistical Techniques

We discussed the following three important statistical techniques:

- How to compute a confidence interval for the population mean μ

Statistical Techniques

We discussed the following three important statistical techniques:

- How to compute a confidence interval for the population mean μ
- How to test a hypothesis based on a sample

Statistical Techniques

We discussed the following three important statistical techniques:

- How to compute a confidence interval for the population mean μ
- How to test a hypothesis based on a sample
- How to compare the means of two samples

Statistical Techniques

We also discussed some special cases for each statistical technique:

- How to compute a confidence interval for the population mean μ
  - When the population standard deviation σ is known
  - When the population standard deviation σ is unknown
  - For a proportion
- How to test a hypothesis based on a sample
  - When the population standard deviation σ is known
  - When the population standard deviation σ is unknown
  - For a proportion
- How to compare the means of two samples
  - When we have paired or dependent samples
  - When the population standard deviation σ is known
  - When the population standard deviation σ is unknown
  - For a proportion

Examples

32 wood ducks are captured and weighed.

The average weight is $\overline{x} = 1.2$ pounds, with a sample standard deviation of $s = 0.4$.

Explain in as much detail as you can which technique you would use to construct a 95% confidence interval for the population mean μ based on this data.

Examples

Solution: We are not told what the population standard deviation, σ is.

Consequently, we would use the technique for unknown σ, which is based on the t-statistic.

Examples

32 wood ducks are captured and weighed.

The average weight is $\overline{x} = 1.2$ pounds.

Data from the previous year indicates that the wood duck population has a mean weight of 1.3 pounds and the standard deviation of a bird's weight is $\sigma = 0.3$.

Does this data suggest that the weight of wood ducks has changed this year?

(explain in as much detail as you can which technique you would use to answer this question)

Examples

Solution: This is a hypothesis test situation. Since we are given σ, we use the technique for known population standard deviation.

Examples

Solution: This is a hypothesis test situation. Since we are given σ, we use the technique for known population standard deviation.

The null hypothesis is usually the claim that nothing has changed.

Consequently, the null hypothesis is that the population mean is 1.3 pounds.

Examples

Solution: This is a hypothesis test situation. Since we are given σ, we use the technique for known population standard deviation.

The null hypothesis is usually the claim that nothing has changed.

Consequently, the null hypothesis is that the population mean is 1.3 pounds.

Since we are concerned only with change, and not with its direction, we would use a two-sided test.

Examples

32 wood ducks are captured and weighed.

The average weight is $\overline{x} = 1.2$ pounds, with a sample standard deviation $s = 0.3$

Data from the previous year indicates that the wood duck population has a mean weight of 1.3 pounds.

Does this data suggest that the weight of wood ducks has *decreased* this year?

(explain in as much detail as you can which technique you would use to answer this question)

Examples

Solution: This is a hypothesis test situation. Since we are **not** given σ, we use the technique for unknown population standard deviation.

Examples

Solution: This is a hypothesis test situation. Since we are **not** given σ, we use the technique for unknown population standard deviation.

The null hypothesis is usually the claim that nothing has changed.

Consequently, the null hypothesis is that the population mean is 1.3 pounds.

Examples

Solution: This is a hypothesis test situation. Since we are **not** given σ, we use the technique for unknown population standard deviation.

The null hypothesis is usually the claim that nothing has changed.

Consequently, the null hypothesis is that the population mean is 1.3 pounds.

Since we are concerned only with a **decrease**, we would use a left-tailed test.
(we want to reject the null hypothesis if the sample mean is small; small values occur to the left).

Examples

32 wood ducks are captured and weighed, and 18 are found to weigh less than 1.3 pounds.

Data from the previous year indicates that 36% weighed less than 1.3 pounds

Does this data suggest that the proportion of wood ducks weighing less than 1.3 pounds is higher this year?

(explain in as much detail as you can which technique you would use to answer this question)

Examples

Solution: This is a hypothesis test on a proportion.

Examples

Solution: This is a hypothesis test on a proportion.

The null hypothesis is usually the claim that nothing has changed.

Consequently, the null hypothesis is that the population proportion is 36%.

Examples

Solution: This is a hypothesis test on a proportion.

The null hypothesis is usually the claim that nothing has changed.

Consequently, the null hypothesis is that the population proportion is 36%.

Since we are concerned only with a **increase**, we would use a right-tailed test.
(we want to reject the null hypothesis if the sample proportion is large; large values occur to the right).

Examples

32 wood ducks are captured and weighed in Rhode Island, and are found to have an average weight of $\overline{x}_1 = 1.2$ pounds with a sample standard deviation $s_1 = 0.4$. A sample of 40 wood ducks is captured in Massachusetts and found to have $\overline{x}_2 = 1.3$ pounds with a sample standard deviation of $s_2 = 0.35$.

Does this data suggest that the mean weight of wood ducks in Rhode Island is different from that in Massachusetts?

(explain in as much detail as you can which technique you would use to answer this question)

Examples

Solution: This is a comparison of two means. Since we are **not** given σ, we use the technique for unknown population standard deviations.

Examples

Solution: This is a comparison of two means. Since we are **not** given σ, we use the technique for unknown population standard deviations.

The null hypothesis in two sample problems is usually that the means of the two populations are the same.

Consequently, the null hypothesis is that the difference in population means is zero.

Examples

Solution: This is a comparison of two means. Since we are **not** given σ, we use the technique for unknown population standard deviations.

The null hypothesis in two sample problems is usually that the means of the two populations are the same.

Consequently, the null hypothesis is that the difference in population means is zero.

Since we are concerned only with a **change**, we would use a two-tailed test.