Sampling Distributions and Confidence Intervals

Chapter 8 and 11, Miller and Miller

November 20, 2015

Big Concepts of the Day

- Population and Parameters

Big Concepts of the Day

- Population and Parameters
- Sample and Statistics

Big Concepts of the Day

- Population and Parameters
- Sample and Statistics
- Grand Question: What do sample statistics tell us about population parameters?

Big Concepts of the Day

- Population and Parameters
- Sample and Statistics
- Grand Question: What do sample statistics tell us about population parameters?
- Key Ingredient: Random Sample

Sample

- X_1, X_2, ... X_n constitutes a random sample of n observations.

Sample

- X_1, X_2, ... X_n constitutes a random sample of n observations.
- Sample mean is given by

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

Sample

- X_1, X_2, ... X_n constitutes a random sample of n observations.
- Sample mean is given by

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

- Sample variance is given by

$$S^2 = \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n-1}$$


WHAT CAN THESE FISH...
...TELL US ABOUT THOSE FISH?

Sampling Distribution

Estimator Types

- Point Estimators: More next week
- Interval Estimators: This week

Progress Map

Statistic	Parameter	Assumptions	Method

Sample Properties

If X_1, X_2, ..., X_n constitutes a random sample from an infinite population with the mean μ and variance σ^2, then

$$E(\bar{X}) = \mu, \qquad var(\bar{X}) = \frac{\sigma^2}{n}$$

Sample Properties

If X_1, X_2, ..., X_n constitutes a random sample from an infinite population with the mean μ and variance σ^2, then

$$E(\bar{X}) = \mu, \qquad var(\bar{X}) = \frac{\sigma^2}{n}$$

(Fun) Proof

Law of Large Numbers

For any positive constant c, the probability that $\bar{X}$ will take on a value between $\mu - c$ and $\mu + c$ is at least

$$1 - \frac{\sigma^2}{nc^2}$$

Law of Large Numbers

For any positive constant c, the probability that $\bar{X}$ will take on a value between $\mu - c$ and $\mu + c$ is at least

$$1 - \frac{\sigma^2}{nc^2}$$

Less practical to use. Primarily of theoretical interest.

Central Limit Theorem

If X_1, X_2, ..., X_n constitutes a random sample from an infinite population with the mean μ, the variance σ^2 and the moment-generating function $M_X(t)$, then the limiting distribution of

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

as $n \to \infty$ is the standard normal distribution.

Central Limit Theorem

If X_1, X_2, ..., X_n constitutes a random sample from an infinite population with the mean μ, the variance σ^2 and the moment-generating function $M_X(t)$, then the limiting distribution of

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

as $n \to \infty$ is the standard normal distribution.
Very important. n should be greater than 30.

Central Limit Theorem

If X_1, X_2, ..., X_n constitutes a random sample from an infinite population with the mean μ, the variance σ^2 and the moment-generating function $M_X(t)$, then the limiting distribution of

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

as $n \to \infty$ is the standard normal distribution.
Very important. n should be greater than 30.
For large samples, σ^2 can be substituted by s^2.

Example 1

For the population of farm workers in New Zealand, suppose that weekly income has a distribution that is skewed right with a mean of \$500 (N.Z. dollars) and a standard deviation of \$160. A survey of 100 farm workers is taken, including information on their weekly income.

- What are the mean and standard error (or margin of error) of the sampling distribution of $\bar{x}$?

Example 1

For the population of farm workers in New Zealand, suppose that weekly income has a distribution that is skewed right with a mean of \$500 (N.Z. dollars) and a standard deviation of \$160. A survey of 100 farm workers is taken, including information on their weekly income.

- What are the mean and standard error (or margin of error) of the sampling distribution of $\bar{x}$?
- What is the probability that the mean weekly income of these 100 workers is less than \$448?

Example 1

For the population of farm workers in New Zealand, suppose that weekly income has a distribution that is skewed right with a mean of \$500 (N.Z. dollars) and a standard deviation of \$160. A survey of 100 farm workers is taken, including information on their weekly income.

- What are the mean and standard error (or margin of error) of the sampling distribution of $\bar{x}$?
- What is the probability that the mean weekly income of these 100 workers is less than \$448?
- What is the probability that the mean weekly income of these 100 workers is between \$480 and \$520?

Example 2

The scores on a general test have mean 450 and standard deviation 50. It is highly desirable to score over 480 on this exam. In one location 45 people sign up to take the exam. The average score of these 45 people exceeds 490. Is this odd? Should the test center investigate? Answer on the basis of the CLT.

Estimation of Population Mean

- You do not know population μ. How to guess it?

Estimation of Population Mean

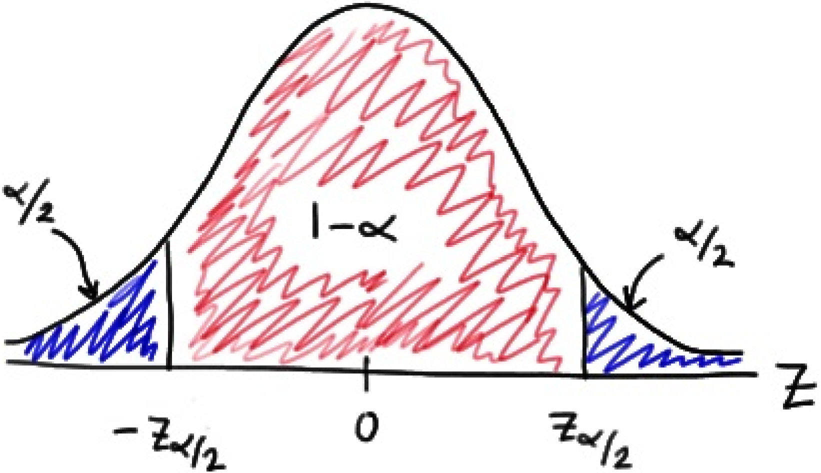
- You do not know population μ. How to guess it?
- Suppose you know the population variance, σ^2

Estimation of Population Mean

- You do not know population μ. How to guess it?
- Suppose you know the population variance, σ^2
- You collect a sample of size $n > 30$, then use sample mean ($\bar{X}$) and Central Limit Theorem.

Estimation of Population Mean

- You do not know population μ. How to guess it?
- Suppose you know the population variance, σ^2
- You collect a sample of size $n > 30$, then use sample mean ($\bar{X}$) and Central Limit Theorem.
- $$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$$
- Assuming with a $(1 - \alpha) * 100$ percent confidence that you have picked a random sample and this sample is not a weird/biased sample, we can say that the probability of your particular sample mean being in proximity of the population mean is $(1 - \alpha)$.



α/2
1−α
α/2
Z
−Zα/2
0
Zα/2

Estimation of Population Mean

$$P\left(-z_{\alpha/2} \leq z \leq z_{\alpha/2}\right) = 1 - \alpha$$

Estimation of Population Mean

$$P\left(-z_{\alpha/2} \leq z \leq z_{\alpha/2}\right) = 1 - \alpha$$

$$P\left(-z_{\alpha/2} \leq \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \leq z_{\alpha/2}\right) = 1 - \alpha$$

$$P\left(-z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \bar{x} - \mu \leq z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

$$P\left(-\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq -\mu \leq -\bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

$$P\left(\bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \geq \mu \geq \bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

$$P\left(\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

Example 3

A questionnaire of spending habits was given to a random sample of college students. Each student was asked to record and report the amount of money they spent on textbooks in a semester. The sample of 130 students resulted in an average of \$422 with standard deviation of \$57.

- Give a 90% confidence interval for the mean amount of money spent by college students on textbooks.

Example 3

A questionnaire of spending habits was given to a random sample of college students. Each student was asked to record and report the amount of money they spent on textbooks in a semester. The sample of 130 students resulted in an average of \$422 with standard deviation of \$57.

- Give a 90% confidence interval for the mean amount of money spent by college students on textbooks.
- Is it true that 90% of the students spent the amount of money found in the interval from previous part? Explain your answer.

Example 3

A questionnaire of spending habits was given to a random sample of college students. Each student was asked to record and report the amount of money they spent on textbooks in a semester. The sample of 130 students resulted in an average of \$422 with standard deviation of \$57.

- Give a 90% confidence interval for the mean amount of money spent by college students on textbooks.
- Is it true that 90% of the students spent the amount of money found in the interval from previous part? Explain your answer.
- What is the margin of error for the 90% confidence interval?

Example 3

A questionnaire of spending habits was given to a random sample of college students. Each student was asked to record and report the amount of money they spent on textbooks in a semester. The sample of 130 students resulted in an average of \$422 with standard deviation of \$57.

- Give a 90% confidence interval for the mean amount of money spent by college students on textbooks.
- Is it true that 90% of the students spent the amount of money found in the interval from previous part? Explain your answer.
- What is the margin of error for the 90% confidence interval?
- How many students should you sample if you want a margin of error of \$5 for a 90% confidence interval?

Example 4

A sample of 30 Stats students yields the sample mean of 82.83. Assume that the population standard deviation is 10.

- Find the 90% confidence interval for the mean score μ for Stats students.

Example 4

A sample of 30 Stats students yields the sample mean of 82.83. Assume that the population standard deviation is 10.

- Find the 90% confidence interval for the mean score μ for Stats students.
- Find the 95% confidence interval.

Example 4

A sample of 30 Stats students yields the sample mean of 82.83. Assume that the population standard deviation is 10.

- Find the 90% confidence interval for the mean score μ for Stats students.
- Find the 95% confidence interval.
- Find the 99% confidence interval.

Example 4

A sample of 30 Stats students yields the sample mean of 82.83. Assume that the population standard deviation is 10.

- Find the 90% confidence interval for the mean score μ for Stats students.
- Find the 95% confidence interval.
- Find the 99% confidence interval.
- How do the margins of error in the previous three questions change as the confidence level increases? Why?

Another Theorem (Z Thm)

If $\bar{X}$ is the mean of a random sample of size n from a normal population with the mean μ and the variance σ^2, its sampling distribution is a normal distribution with the mean μ and the variance σ^2/n.

Another Theorem (Z Thm)

If $\bar{X}$ is the mean of a random sample of size n from a normal population with the mean μ and the variance σ^2, its sampling distribution is a normal distribution with the mean μ and the variance σ^2/n.
Size of n not a concern.

Example 5

Suppose that you have a sample of 10 values from a population with mean $\mu = 50$ and with standard deviation $\sigma = 8$.

- What is the probability that the sample mean will be in the interval (49 51)?

Example 5

Suppose that you have a sample of 10 values from a population with mean $\mu = 50$ and with standard deviation $\sigma = 8$.

- What is the probability that the sample mean will be in the interval (49 51)?
- Suppose the population is normal. What is the probability that the sample mean will be in the interval (49 51)?

Example 5

Suppose that you have a sample of 10 values from a population with mean $\mu = 50$ and with standard deviation $\sigma = 8$.

- What is the probability that the sample mean will be in the interval (49 51)?
- Suppose the population is normal. What is the probability that the sample mean will be in the interval (49 51)?
- Suppose the population is normal. Give an interval that covers the middle 95% of the distribution of the sample mean.

Difference of Normal RVs

If there are two independent random variables such that $X_1 \sim N(\mu_1, \sigma_1^2)$ and $X_2 \sim N(\mu_2, \sigma_2^2)$, the difference of the two random variables is a random variable (lets denote by X) has the following property

Difference of Normal RVs

If there are two independent random variables such that $X_1 \sim N(\mu_1, \sigma_1^2)$ and $X_2 \sim N(\mu_2, \sigma_2^2)$, the difference of the two random variables is a random variable (lets denote by X) has the following property (guess)

Difference of Normal RVs

If there are two independent random variables such that $X_1 \sim N(\mu_1, \sigma_1^2)$ and $X_2 \sim N(\mu_2, \sigma_2^2)$, the difference of the two random variables is a random variable (lets denote by X) has the following property (guess)

$$X \equiv X_1 - X_2 \sim N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2)$$

Difference of Means

For independent random samples from normal populations, $\bar{X}_1 \sim N(\mu_1, \sigma_1^2/n_1)$ and $\bar{X}_2 \sim N(\mu_2, \sigma_2^2/n_2)$, then

$$Z = \bar{X}_1 - \bar{X}_2 \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$$

and hence

$$(\bar{x}_1 - \bar{x}_2) - z_{\alpha/2} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + z_{\alpha/2} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Difference of Means

For independent random samples from normal populations, $\bar{X}_1 \sim N(\mu_1, \sigma_1^2/n_1)$ and $\bar{X}_2 \sim N(\mu_2, \sigma_2^2/n_2)$, then

$$Z = \bar{X}_1 - \bar{X}_2 \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$$

and hence

$$(\bar{x}_1 - \bar{x}_2) - z_{\alpha/2} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + z_{\alpha/2} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

By virtue of central limit theorem, this confidence interval formula can also be used for independent random samples from non-normal populations with $n_1 \geq 30$ and $n_2 \geq 30$.

Example 6

Construct a 94% confidence interval for the difference between stats scores of boys and girls, given that random sample of 6 girls have an average score of 63 and the average score of 6 boys is 48. The populations scores are normally distributed with $\sigma_g = 8$ and $\sigma_b = 10$.

Example 6

Construct a 94% confidence interval for the difference between stats scores of boys and girls, given that random sample of 6 girls have an average score of 63 and the average score of 6 boys is 48. The populations scores are normally distributed with $\sigma_g = 8$ and $\sigma_b = 10$.
Based on true events surrounding us.

CLT for estimating Proportions

1. We want to estimate the proportion of people in the US who wear corrective lenses.
2. We want to estimate what fraction of Delhi citizens text while driving?
3. We want to estimate what percentage of farmers favor Land Acquisition Bill?
4. We want to estimate what fraction of voters will vote for BJP, Congress or Left?

Method for Estimating Proportions

1. You are estimating the population proportion, p (or θ).

Method for Estimating Proportions

1. You are estimating the population proportion, p (or θ).
2. All estimation done here is based on the fact that the normal can be used to approximate the binomial distribution when *np* and *nq* are both at least 5, p is the probability of success on a single trial from the binomial experiments.

Method for Estimating Proportions

1. You are estimating the population proportion, p (or θ).
2. All estimation done here is based on the fact that the normal can be used to approximate the binomial distribution when np and nq are both at least 5, p is the probability of success on a single trial from the binomial experiments.
3. The best point estimate for p is $\hat{p}$, the sample proportion is $\hat{p} = x/n$, where x is the number of successes observed in sample of size n.

Method for Estimating Proportions

1. You are estimating the population proportion, p (or θ).
2. All estimation done here is based on the fact that the normal can be used to approximate the binomial distribution when np and nq are both at least 5, p is the probability of success on a single trial from the binomial experiments.
3. The best point estimate for p is $\hat{p}$, the sample proportion is $\hat{p} = x/n$, where x is the number of successes observed in sample of size n.
4. $$Z = \frac{x - np}{\sqrt{npq}} = \frac{n\hat{p} - np}{\sqrt{npq}} \sim N(0, 1)$$

Method for Estimating Proportions

1. You are estimating the population proportion, p (or θ).
2. All estimation done here is based on the fact that the normal can be used to approximate the binomial distribution when np and nq are both at least 5, p is the probability of success on a single trial from the binomial experiments.
3. The best point estimate for p is $\hat{p}$, the sample proportion is $\hat{p} = x/n$, where x is the number of successes observed in sample of size n.
4. $$Z = \frac{x - np}{\sqrt{npq}} = \frac{n\hat{p} - np}{\sqrt{npq}} \sim N(0,1)$$
5. $$\frac{\hat{p} - p}{\sqrt{\hat{p}(1-\hat{p})/n}} \sim N(0,1)$$

Z Thm for Proportion

$$P\left(\hat{p} - \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} * z_{\alpha/2} \leq p \leq \hat{p} + \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} * z_{\alpha/2}\right) = 1 - \alpha$$

Example 7

Random sample of 935 Americans were surveyed "Do you think there is intelligent life on other planets?" 60% of the sample said "yes" Calculate a 98% confidence interval of the population proportion.

Example 7

Random sample of 935 Americans were surveyed “Do you think there is intelligent life on other planets?” 60% of the sample said “yes” Calculate a 98% confidence interval of the population proportion.

$$\hat{p} = 0.6, \qquad n = 935$$

Example 7

Random sample of 935 Americans were surveyed "Do you think there is intelligent life on other planets?" 60% of the sample said "yes" Calculate a 98% confidence interval of the population proportion.

$$\hat{p} = 0.6, \qquad n = 935$$

$$\left(\hat{p} - \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} * z_{.01}, \quad \hat{p} + \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} * z_{.01} \right)$$

Difference of Proportions

For independent random binomial populations, $X_1 \sim B(n_1, \theta_1)$ and $X_2 \sim B(n_2, \theta_2)$, and

$$\hat{\theta}_1 = \frac{x_1}{n_1}, \qquad \hat{\theta}_2 = \frac{x_2}{n_2}$$

, then

$$Z = \hat{\Theta}_1 - \hat{\Theta}_2 \sim N\left(\theta_1 - \theta_2, \frac{\theta_1(1-\theta_1)}{n_1} + \frac{\theta_2(1-\theta_2)}{n_2}\right)$$

and hence

$$(\hat{\theta}_1 - \hat{\theta}_2) - z_{\alpha/2} \cdot \sqrt{\frac{\hat{\theta}_1(1-\hat{\theta}_1)}{n_1} + \frac{\hat{\theta}_2(1-\hat{\theta}_2)}{n_2}} < \theta_1 - \theta_2$$
$$< (\hat{\theta}_1 - \hat{\theta}_2) + z_{\alpha/2} \cdot \sqrt{\frac{\hat{\theta}_1(1-\hat{\theta}_1)}{n_1} + \frac{\hat{\theta}_2(1-\hat{\theta}_2)}{n_2}}$$

Example 8

In a random sample of visitors to Taj Mahal, 84 of 250 Indians and 156 of 250 foreigners bought souvenirs. Construct a 95% confidence interval for the difference between the true proportion of Indians and foreigners who buy souvenirs at Taj Mahal.

Chi-Square Distribution and Variance

- Gujrathi Notes.

Chi-Square Distribution and Variance

- Gujrathi Notes.
- If $\bar{X}$ and S^2 are the mean and the variance of a random sample of size n from a normal population with the mean μ and the standard deviation σ then
  (a) $\bar{X}$ and S^2 are independent;
  (b) the random variable $(n-1)S^2/\sigma^2$ has a chi-square distribution with $n-1$ degrees of freedom.

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

Example 9

The claim that the variance of a normal population is $\sigma^2 = 25$ is to be rejected if the variance of a random sample of size 16 exceeds 54.668 or is less than 12.102. What is the probability that this claim will be rejected even though $\sigma^2 = 25$?

Estimating Variance

If s^2 is the variance of a random sample of size n from a normal population, the

$$\frac{(n-1)s^2}{\chi^2_{\alpha/2,n-1}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{1-\alpha/2,n-1}}$$

is a $(1-\alpha)100\%$ confidence interval for σ^2.

Example 10

The length of the skulls of 10 fossil skeletons of an extinct species of bird has a mean of 5.68 cm and a standard deviation of 0.29 cm. Assuming that such a measurement is normally distributed, construct a 95% confidence interval for the true variance of the skull length of the given species of bird.

Progress Map I

Sample Statistic	Parameter	Assumptions	Test Statistic	Method
$\bar{X}$	μ	Large sample ($n \geq 30$). s^2 can also substitute for σ^2	$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$	Central Limit Theorem
$\bar{X}$	μ	Normal population. σ^2 known	$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$	Z Theorem
$\hat{p}$ or $\hat{\theta}$	Pop. Prop. (p or θ)	$n \geq 30$, $n\hat{p} > 5$ and $n(1-\hat{p}) > 5$	$\frac{\hat{p}-p}{\hat{p}(1-\hat{p})/\sqrt{n}}$	Z Theorem for Prop.
$\bar{X}_1 - \bar{X}_2$	$\mu_1 - \mu_2$	Normal pop. and σ^2 known, or Large Sample ($n \geq 30$)	See Slide 20	Z Theorem or CLT
$\hat{\theta}_1 - \hat{\theta}_2$	Pop. Prop. Diff. ($\theta_1 - \theta_2$)	n_1, $n_2 \geq 30$	See Slide 26	Z Theorem for Proportions
s^2	σ^2	Normal Population	See Slide 30	Chi-Square

T Distribution and Sample Mean

- Gujrathi Notes.

T Distribution and Sample Mean

- Gujrathi Notes.
- If $\bar{X}$ and S^2 are the mean and the variance of a random sample of size n from a normal population with the mean μ and the standard deviation σ then
$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t(n-1)$$
has a t distribution with $n-1$ degrees of freedom.

Example 11

A random sample of size $n = 12$ from a normal population has the mean $\bar{x} = 27.8$ and the variance $s^2 = 3.24$. Can we say that the given information supports the claim that the mean of the population is $\mu = 28.5$?

Means of Small Normal Samples

If $\bar{x}$ and s are the values of the mean and the standard deviation of a random sample of size n from a normal population then

$$\bar{x} - t_{\alpha/2,n-1} \cdot \frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{\alpha/2,n-1} \cdot \frac{s}{\sqrt{n}}$$

is a $(1-\alpha)100\%$ confidence interval for the mean of the population.

Example 12

The length of the skulls of 10 fossil skeletons of an extinct species of bird has a mean of 5.68 cm and a standard deviation of 0.29 cm. Assuming that such a measurement is normally distributed, construct a 95% confidence interval for the mean length of the skull length of the given species of bird.

Difference of Means, Small Normal Population

If $\bar{x}_1$, $\bar{x}_2$, s_1 and s_2 are the values of the means and the standard deviations of independent random samples of sizes n_1 and n_2 from normal populations with equal variances, then

$$(\bar{x}_1 - \bar{x}_2) - t_{\alpha/2, n_1+n_2-2} \cdot s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} < \mu_1 - \mu_2$$
$$< (\bar{x}_1 - \bar{x}_2) + t_{\alpha/2, n_1+n_2-2} \cdot s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

where

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

is a $(1 - \alpha)100\%$ confidence interval for the difference between the two population means.

Example 13

The heat-producing capacities of coal from two mines are

	Mine A	Mine B
Sample Mean	8260	7930
Sample Variance	63450	42650

Assuming that the data constitute independent random samples from normal populations with equal variances, construct a 99% confidence interval for the difference between the true average heat-producing capacities of coal from the two mines.

F Distribution and Variance of two Samples

- Gujrathi Notes.

F Distribution and Variance of two Samples

- Gujrathi Notes.
- If S_1^2 and S_2^2 are the variances of independent random samples of sizes n_1 and n_2 from a <u>normal</u> population with the variances σ_1^2 and σ_2^2, then

$$F = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F(n_1 - 1, n_2 - 1)$$

  has a F distribution with $n_1 - 1$ and $n_2 - 1$ degrees of freedom.

Example 14

If S_1^2 and S_2^2 are the standard deviations of independent random samples of sizes $n_1 = 61$ and $n_2 = 31$ from normal populations with $\sigma_1^2 = 12$ and $\sigma_2^2 = 18$, find $P(S_1^2/S_2^2 > 1.16)$.

Estimating ratio of variances

If s_1^2 and s_2^2 are the values of the variances of independent random samples of sizes n_1 and n_2 from normal populations, then

$$\frac{s_1^2}{s_2^2} \cdot f_{1-\alpha/2, n_2-1, n_1-1} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{s_1^2}{s_2^2} \cdot f_{\alpha/2, n_2-1, n_1-1}$$

is a $(1-\alpha)100\%$ confidence interval for σ_1^2/σ_2^2.

Estimating ratio of variances

If s_1^2 and s_2^2 are the values of the variances of independent random samples of sizes n_1 and n_2 from normal populations, then

$$\frac{s_1^2}{s_2^2} \cdot f_{1-\alpha/2, n_2-1, n_1-1} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{s_1^2}{s_2^2} \cdot f_{\alpha/2, n_2-1, n_1-1}$$

is a $(1-\alpha)100\%$ confidence interval for σ_1^2/σ_2^2.
Note:

$$f_{1-\alpha/2, n_2-1, n_1-1} = \frac{1}{f_{\alpha/2, n_1-1, n_2-1}}$$

Example 15

The heat-producing capacities of coal from two mines are

	Mine A	Mine B
Sample Mean	8260	7930
Sample Variance	63450	42650

Construct a 90% confidence interval for the ratio of the variances of the two populations samples.

Progress Map II

Sample Statistic	Parameter	Assumptions	Test Statistic	Method
$\bar{X}$	μ	Small normal sample ($n \leq 30$)	$\frac{\bar{X}-\mu}{s/\sqrt{n}}$, Slide 35	T Test
$\bar{X}_1 - \bar{X}_2$	μ	Small normal samples, common σ^2	Slide 37	T Test
$\frac{s_1^2}{s_2^2}$	$\frac{\sigma_1^2}{\sigma_2^2}$	Samples from normal populations	See Slide 41	F test