Research Methods for Social Sciences:

Point and Interval Estimations; Hypothesis Tests

Alex Teytelboym & Earth Sivakul

@t8el

Hilary 2013

Outline

1. Point and Interval Estimations; Hypothesis tests
    - Point estimation
    - Interval estimation
    - Hypothesis test: large samples
    - Hypothesis test: small samples

Two types of estimators

- Point estimator:
  - Consists of a single number, calculated from the data, that is the best single guess for the parameter

Example

In a survey, 1500 people were asked 'Do you support the introduction of the energy bill?' The point estimate for the proportion of all Americans who support this Bill is 0.63

Two types of estimators

- Point estimator:
  - Consists of a single number, calculated from the data, that is the best single guess for the parameter

Example

In a survey, 1500 people were asked 'Do you support the introduction of the energy bill?' The point estimate for the proportion of all Americans who support this Bill is 0.63

Two types of estimators

- Interval estimator:
  - Consists of a range of numbers around the point estimate, within which the parameter is believed to fall.
  - It helps to gauge the probable accuracy of a sample point estimate.

Example

In a survey, 1500 people were asked 'Do you support the introduction of the energy bill?' An interval estimate at 95% confidence level predicts that the population proportion responding 'yes' falls between 0.59 and 0.67.

Two types of estimators

- Interval estimator:
  - Consists of a range of numbers around the point estimate, within which the parameter is believed to fall.
  - It helps to gauge the probable accuracy of a sample point estimate.

Example

In a survey, 1500 people were asked ‘Do you support the introduction of the energy bill?’ An interval estimate at 95% confidence level predicts that the population proportion responding ‘yes’ falls between 0.59 and 0.67.

Two types of estimators

- Interval estimator:
  - Consists of a range of numbers around the point estimate, within which the parameter is believed to fall.
  - It helps to gauge the probable accuracy of a sample point estimate.

Example

In a survey, 1500 people were asked 'Do you support the introduction of the energy bill?' An interval estimate at 95% confidence level predicts that the population proportion responding 'yes' falls between 0.59 and 0.67.

Interval estimation

- Confidence interval estimate: range of numbers around the point estimate, within which the parameter is believed to fall with a given probability (confidence level).
- The confidence gives the probable accuracy of the estimate.
- The width of confidence interval depends on standard error (SE) of the estimator, for the given sample size.

Interval estimation

- Confidence interval estimate: range of numbers around the point estimate, within which the parameter is believed to fall with a given probability (confidence level).
- The confidence gives the probable accuracy of the estimate.
- The width of confidence interval depends on standard error (SE) of the estimator, for the given sample size.

Interval estimation

- Confidence interval estimate: range of numbers around the point estimate, within which the parameter is believed to fall with a given probability (confidence level).
- The confidence gives the probable accuracy of the estimate.
- The width of confidence interval depends on standard error (SE) of the estimator, for the given sample size.

Interval estimation

- The width of confidence interval depends on standard error (SE) of the estimator, for the given sample size.
- The smaller the SE, the narrower confidence interval, and the more accurate (efficient) the estimate.
- When a sampling distribution is approximately normal, we construct a confidence interval by adding to and subtracting from the point estimate some multiple (a z-score) of its standard error.
- The most commonly utilised confidence levels are 95%, 90%, and 99%.

Interval estimation

- The width of confidence interval depends on standard error (SE) of the estimator, for the given sample size.
- The smaller the SE, the narrower confidence interval, and the more accurate (efficient) the estimate.
- When a sampling distribution is approximately normal, we construct a confidence interval by adding to and subtracting from the point estimate some multiple (a z-score) of its standard error.
- The most commonly utilised confidence levels are 95%, 90%, and 99%.

Interval estimation

- The width of confidence interval depends on standard error (SE) of the estimator, for the given sample size.
- The smaller the SE, the narrower confidence interval, and the more accurate (efficient) the estimate.
- When a sampling distribution is approximately normal, we construct a confidence interval by adding to and subtracting from the point estimate some multiple (a z-score) of its standard error.
- The most commonly utilised confidence levels are 95%, 90%, and 99%.

Interval estimation

- The width of confidence interval depends on standard error (SE) of the estimator, for the given sample size.
- The smaller the SE, the narrower confidence interval, and the more accurate (efficient) the estimate.
- When a sampling distribution is approximately normal, we construct a confidence interval by adding to and subtracting from the point estimate some multiple (a z-score) of its standard error.
- The most commonly utilised confidence levels are 95%, 90%, and 99%.

Confidence interval for the mean (large sample)

- The confidence interval for the mean is calculated in the following way:

$$\mu = \overline{Y} \pm z_{\alpha/2} \times SE = \overline{Y} \pm z_{\alpha/2} \times \frac{s}{\sqrt{n}}$$

  where

  - $\overline{Y}$ is the sample average
  - α is error probability or significance level of the interval . It is the probability that the interval does not contain the μ.
  - SE is the standard error (i.e. the standard deviation of the sampling distribution of the population mean)
  - $z_{\alpha/2}$: standardised value (z-score) that leaves $\alpha/2$ probability on one side, and $\alpha/2$ on the other side

Confidence interval for the mean (large sample)

- We could rewrite this as:

$$P\left[\overline{Y} - z_{\alpha/2} \times \frac{s}{\sqrt{n}} < \mu < \overline{Y} + z_{\alpha/2} \times \frac{s}{\sqrt{n}}\right] = 1 - \alpha$$

where

  - $\overline{Y}$ is the sample average
  - s is the standard deviation of the sample
  - α is error probability or significance level of the interval . It is the probability that the interval does not contain the μ.
  - $z_{\alpha/2}$: standardised value (z-score) that leaves $\alpha/2$ probability on one side, and $\alpha/2$ on the other side

z-scores

Table: z-scores

Confidence level $(1-\alpha)$	Error probability/ significance level α	$\alpha/2$	$z_{\alpha/2}$	Confidence interval
90%	10%	0.05	1.64	$\mu = \overline{Y} \pm 1.65 \times SE$
95%	5%	0.025	1.96	$\mu = \overline{Y} \pm 1.96 \times SE$
99%	1%	0.005	2.58	$\mu = \overline{Y} \pm 2.58 \times SE$

z-scores

- Interpretation: a 95% confidence interval : if an infinite number of confidence intervals are constructed from random samples of a given size, 95% of them would contain the population mean. In other words, we can be 95% confident that this interval contains the true μ.
- The larger the interval, the less accurate is the estimation. For given μ and σ, width of confidence interval estimates of μ can be controlled by manipulating. . .
- . . .sample size: increasing the sample size, decreases the $SE \Rightarrow$ decreases the confidence interval width, more accurate or precise is the estimation.
- As sample size goes to infinity, SE tends to zero, the confidence interval goes to zero and the sample average approaches the population mean.

z-scores

- Interpretation: a 95% confidence interval : if an infinite number of confidence intervals are constructed from random samples of a given size, 95% of them would contain the population mean. In other words, we can be 95% confident that this interval contains the true μ.
- The larger the interval, the less accurate is the estimation. For given μ and σ, width of confidence interval estimates of μ can be controlled by manipulating...
- ...sample size: increasing the sample size, decreases the $SE \Rightarrow$ decreases the confidence interval width, more accurate or precise is the estimation.
- As sample size goes to infinity, SE tends to zero, the confidence interval goes to zero and the sample average approaches the population mean.

z-scores

- Interpretation: a 95% confidence interval : if an infinite number of confidence intervals are constructed from random samples of a given size, 95% of them would contain the population mean. In other words, we can be 95% confident that this interval contains the true μ.
- The larger the interval, the less accurate is the estimation. For given μ and σ, width of confidence interval estimates of μ can be controlled by manipulating...
- ...sample size: increasing the sample size, decreases the $SE \Rightarrow$ decreases the confidence interval width, more accurate or precise is the estimation.
- As sample size goes to infinity, SE tends to zero, the confidence interval goes to zero and the sample average approaches the population mean.

z-scores

- Interpretation: a 95% confidence interval : if an infinite number of confidence intervals are constructed from random samples of a given size, 95% of them would contain the population mean. In other words, we can be 95% confident that this interval contains the true μ.
- The larger the interval, the less accurate is the estimation. For given μ and σ, width of confidence interval estimates of μ can be controlled by manipulating...
- ...sample size: increasing the sample size, decreases the *SE* $\Rightarrow$ decreases the confidence interval width, more accurate or precise is the estimation.
- As sample size goes to infinity, *SE* tends to zero, the confidence interval goes to zero and the sample average approaches the population mean.

Confidence intervals

Example

- Life expectancy at birth (data for 2003, UNDP)
- Sample size is 177 countries

Table: Life expectancy at birth

	Life expectancy at birth (years)
Mean	65.81
Standard deviation	12.276

Confidence intervals

Example

- Life expectancy at birth (data for 2003, UNDP)
- $\overline{Y} = 65.8$, $s = 12.276$, $n = 177$
- 99% confidence interval of life expectancy (LE) in the population is:

$$65.8 \pm 2.58 \times \frac{12.276}{\sqrt{177}} = 65.8 \pm 2.38 = [63.42, 68.18]$$

- If we decrease the confidence level: 90% confidence interval for LE in the population is:

$$65.8 \pm 1.65 \times \frac{12.276}{\sqrt{177}} = 65.8 \pm 1.52 = [64.28, 67.32]$$

Confidence intervals

Example

- Life expectancy at birth (data for 2003, UNDP)
- $\overline{Y} = 65.8$, $s = 12.276$, $n = 177$
- 99% confidence interval of life expectancy (LE) in the population is:

$$65.8 \pm 2.58 \times \frac{12.276}{\sqrt{177}} = 65.8 \pm 2.38 = [63.42, 68.18]$$

- If we increase the sample size to $n = 1770$, 99% conf interval for LE in the population is:

$$65.8 \pm 2.58 \times \frac{12.276}{\sqrt{1770}} = 65.8 \pm 0.752 = [65.05, 66.55]$$

Hypothesis Test

- Statistical hypothesis: statement about a characteristic (parameter) of a population.
- A hypothesis test (or significance test) uses sample data to test validity of hypotheses.
- Five elements of the hypothesis testing
  1. Assumption
  2. Hypothesis
  3. Test statistic
  4. *P*-value
  5. Conclusion

Hypothesis Test

- Statistical hypothesis: statement about a characteristic (parameter) of a population.
- A hypothesis test (or significance test) uses sample data to test validity of hypotheses.
- Five elements of the hypothesis testing
  1. Assumption
  2. Hypothesis
  3. Test statistic
  4. P-value
  5. Conclusion

Hypothesis Test

- Statistical hypothesis: statement about a characteristic (parameter) of a population.
- A hypothesis test (or significance test) uses sample data to test validity of hypotheses.
- Five elements of the hypothesis testing
  1. Assumption
  2. Hypothesis
  3. Test statistic
  4. P-value
  5. Conclusion

Five elements of hypothesis testing

- Assumptions
  - type of data
  - population distribution
  - sampling technique (simple random)
  - sample size
- Hypothesis:
  - Null hypothesis H_0: $\mu = \mu_0$ (the hypothesis that is directly tested)
  - Alternative hypothesis H_A: $\mu \neq \mu_0$, contradicts the null (research hypothesis)
  - If H_0 is rejected, H_A is not rejected.
  - Hypothesis tests investigate whether sample data are inconsistent with H_0 so that H_A is true.

Five elements of hypothesis testing

- Assumptions
  - type of data
  - population distribution
  - sampling technique (simple random)
  - sample size
- Hypothesis:
  - Null hypothesis H_0: $\mu = \mu_0$ (the hypothesis that is directly tested)
  - Alternative hypothesis H_A: $\mu \neq \mu_0$, contradicts the null (research hypothesis)
  - If H_0 is rejected, H_A is not rejected.
  - Hypothesis tests investigate whether sample data are inconsistent with H_0 so that H_A is true.

Five elements of hypothesis testing

- Test statistic
  - sample statistic to test the null hypothesis
  - for example the *z*-score for means and proportions compares point estimate to null hypothesis in terms of standard errors:

$$z = \frac{\overline{Y} - \mu}{SE}$$

Five elements of hypothesis testing

- *P*-value
  - probability of obtaining the observed test statistic if H_0 is true.
  - The smaller *P*-value, the more strongly the data contradicts the H_0. If the *P*-value $= 0.01$, the probability of obtaining observed statistic if null were true $= 1\%$.
- Conclusion
  - If *P*-value ‘sufficiently’ small, H_0 is rejected in favour of H_A.

Five elements of hypothesis testing

- P-value
  - probability of obtaining the observed test statistic if H_0 is true.
  - The smaller P-value, the more strongly the data contradicts the H_0. If the P-value $= 0.01$, the probability of obtaining observed statistic if null were true $= 1\%$.
- Conclusion
  - If P-value 'sufficiently' small, H_0 is rejected in favour of H_A.

Decision rule

- Decision Rule
  - using a pre-specified defined significance level (α) as a cut-off point to make a decision in terms of rejecting the null or not.
  - choice of α depends on desired caution. Lower $\alpha \Rightarrow$ more caution.
  - The z-score corresponding to the chosen level of α is called critical value.
- Hence:
  - if P-value $\leq \alpha$ (or if $|z| \geq z_\alpha$) $\Rightarrow$ reject H_0
  - if P-value $> \alpha$ (or if $|z| < z_\alpha$)$\Rightarrow$ do not reject H_0

Decision rule

- Decision Rule
  - using a pre-specified defined significance level (α) as a cut-off point to make a decision in terms of rejecting the null or not.
  - choice of α depends on desired caution. Lower $\alpha \Rightarrow$ more caution.
  - The z-score corresponding to the chosen level of α is called critical value.
- Hence:
  - if P-value $\leq \alpha$ (or if $|z| \geq z_\alpha$) $\Rightarrow$ reject H_0
  - if P-value $> \alpha$ (or if $|z| < z_\alpha$)$\Rightarrow$ do not reject H_0

Decision rule

- H_0 and α must be chosen before data are analysed (to avoid biases).
- Never accept the null hypothesis; only fail to reject it.
- But if we reject the null hypothesis, we accept H_A. This gives us a range of possible parameter values.
- Hypothesis significance does not imply practical significance: parameter may be statistically significantly different from H_0, but difference for us as statisticians (or development economists) might be unimportant. This happens a lot.

Decision rule

- H_0 and α must be chosen before data are analysed (to avoid biases).
- Never accept the null hypothesis; only fail to reject it.
- But if we reject the null hypothesis, we accept H_A. This gives us a range of possible parameter values.
- Hypothesis significance does not imply practical significance: parameter may be statistically significantly different from H_0, but difference for us as statisticians (or development economists) might be unimportant. This happens a lot.

Decision rule

- H_0 and α must be chosen before data are analysed (to avoid biases).
- Never accept the null hypothesis; only fail to reject it.
- But if we reject the null hypothesis, we accept H_A. This gives us a range of possible parameter values.
- Hypothesis significance does not imply practical significance: parameter may be statistically significantly different from H_0, but difference for us as statisticians (or development economists) might be unimportant. This happens a lot.

Decision rule

- H_0 and α must be chosen before data are analysed (to avoid biases).
- Never accept the null hypothesis; only fail to reject it.
- But if we reject the null hypothesis, we accept H_A. This gives us a range of possible parameter values.
- Hypothesis significance does not imply practical significance: parameter may be statistically significantly different from H_0, but difference for us as statisticians (or development economists) might be unimportant. This happens a lot.

Hypothesis test for a mean (large samples)

- Assumptions:
  - quantitative variable
  - random sample
  - sample size ≥ 30
  - then we know (by the Central Limit Theorem) that $\overline{Y} \sim N(\mu, \frac{\sigma}{\sqrt{n}})$
- Hypotheses
  - H_0: $\mu = \mu_0$ where μ_0 is some particular value.
  - H_A: $\mu \neq \mu_0$ (two sided test) or $\mu > \mu_0$, $\mu < \mu_0$ (one sided test)

Hypothesis test for a mean (large samples)

- Assumptions:
  - quantitative variable
  - random sample
  - sample size ≥ 30
  - then we know (by the Central Limit Theorem) that $\overline{Y} \sim N(\mu, \frac{\sigma}{\sqrt{n}})$
- Hypotheses
  - H_0: $\mu = \mu_0$ where μ_0 is some particular value.
  - H_A: $\mu \neq \mu_0$ (two sided test) or $\mu > \mu_0$, $\mu < \mu_0$ (one sided test)

Hypothesis test for a mean (large samples)

- Test statistic if the null is true:

$$z = \frac{\overline{Y} - \mu}{SE} \sim N(0, 1)$$

- P-value
  - probability of obtaining the z-score, if H_0 were true. As sample size increases, standard error (SE) falls for a given sample mean, z-score rises and P-value falls.
- Conclusion
  - if the P-value is small, reject the null!

Hypothesis test for a mean (large samples)

- Test statistic if the null is true:

$$z = \frac{\overline{Y} - \mu}{SE} \sim N(0, 1)$$

- *P*-value
  - probability of obtaining the z-score, if H_0 were true. As sample size increases, standard error (SE) falls for a given sample mean, z-score rises and P-value falls.
- Conclusion
  - if the P-value is small, reject the null!

Hypothesis test for a mean (large samples)

- Test statistic if the null is true:

$$z = \frac{\overline{Y} - \mu}{SE} \sim N(0,1)$$

- *P*-value
  - probability of obtaining the z-score, if H_0 were true. As sample size increases, standard error (SE) falls for a given sample mean, z-score rises and P-value falls.
- Conclusion
  - if the P-value is small, reject the null!

Hypothesis test for a mean

Example

- Returning to the example above
- Life expectancy at birth (data for 2003, UNDP)
- An international organisation has launched a programme to improve health care. One of the objectives is to enhance life expectancy to exceed 65 years.
- A random survey has been launched with 177 countries in 2003. The statistics from the sample is as follows. Has this objective of the programme being achieved?

Hypothesis test for a mean

Example

- Life expectancy at birth (data for 2003, UNDP)

Table: Life expectancy at birth

	Life expectancy at birth (years)
Mean	65.81
Standard deviation	12.276

Hypothesis test for a mean

Example

- Life expectancy at birth (data for 2003, UNDP)
- Assumptions are satisfied
  - quantitative variable
  - random sample
  - sample size $177 \geq 30$
- Hypotheses
  - H_0: $\mu = 65$
  - H_A: $\mu > 65$, (one sided test)

Hypothesis test for a mean

Example

- Life expectancy at birth (data for 2003, UNDP)
- Assumptions are satisfied
  - quantitative variable
  - random sample
  - sample size $177 \geq 30$
- Hypotheses
  - H_0: $\mu = 65$
  - H_A: $\mu > 65$, (one sided test)

Hypothesis test for a mean

Example

- Life expectancy at birth (data for 2003, UNDP)
- Test statistic if the null is true:

$$z = \frac{65.8 - 65}{12.276/\sqrt{177}} = 0.87$$

- P-value
  - P-value=0.1922
- Conclusion
  - P-value > 0.10, so cannot reject H_0 that life expentancy has remained the same. Therefore, the objective of increasing life expentancy has not been achieved.

Hypothesis test for a mean

Example

- Life expectancy at birth (data for 2003, UNDP)
- Test statistic if the null is true:

$$z = \frac{65.8 - 65}{12.276/\sqrt{177}} = 0.87$$

- *P*-value
  - *P*-value=0.1922
- Conclusion
  - *P*-value > 0.10, so cannot reject H_0 that life expentancy has remained the same. Therefore, the objective of increasing life expentancy has not been achieved.

Hypothesis test for a mean

Example

- Life expectancy at birth (data for 2003, UNDP)
- Test statistic if the null is true:

$$z = \frac{65.8 - 65}{12.276/\sqrt{177}} = 0.87$$

- *P*-value
  - *P*-value=0.1922
- Conclusion
  - *P*-value > 0.10, so cannot reject H_0 that life expentancy has remained the same. Therefore, the objective of increasing life expentancy has not been achieved.

Hypothesis test for a mean

Example

- Life expectancy at birth (data for 2003, UNDP)
- Test statistic if the null is true:

$$z = \frac{65.8 - 65}{12.276/\sqrt{177}} = 0.87$$

- P-value
  - P-value=0.1922
- Conclusion
  - Note that we could have just as well said: "I want a P-value of at most 0.1 so the z-score has to be greater than 1.64 in order to reject the null. It is not, so I cannot reject the null!

Hypothesis test for a mean

Example

- Life expectancy at birth (data for 2003, UNDP)
- Test statistic if the null is true:

$$z = \frac{65.8 - 65}{12.276/\sqrt{177}} = 0.87$$

- *P*-value
  - *P*-value=0.1922
- Conclusion
  - Note that we could have just as well said: "I want a *P*-value of at most 0.1 so the z-score has to be greater than 1.64 in order to reject the null. It is not, so I cannot reject the null!

Hypothesis test for a mean

Example

- Life expectancy at birth (data for 2003, UNDP)
- Test statistic if the null is true:

$$z = \frac{65.8 - 65}{12.276/\sqrt{177}} = 0.87$$

- *P*-value
  - *P*-value=0.1922
- Conclusion
  - Note that we could have just as well said: "I want a *P*-value of at most 0.1 so the *z*-score has to be greater than 1.64 in order to reject the null. It is not, so I cannot reject the null!

Aside: one-sided and two-sided tests

- In order to work out the P-value for a one-sided test, you simply need to use the z-score appropriately
- A two-sided test works like a confidence interval calculation. Suppose you want a P-value of 0.1 for a two-sided test. You want 5% of the normal distribution to like outside the modulus of the correct z-score. So, look up the closest z-score for $z \leq 0.95$, which is 1.64. So any z-score above 1.64 or below -1.64 will make you reject the null hypothesis at the 90% confidence level.

Types of errors in hypothesis tests

- Type I error: rejecting H_0 when it is true.
- Type II error: not rejecting H_0 when it is false
- The probability of Type I and Type II error are inversely related

Table: Types of errors

	H_0 is false	H_0 is true
Reject H_0	Correct decision	Type I error (w.p. α)
Fail to reject H_0	Type II error	Correct decision

- Statistics joke from Casella and Berger: one ought not forget the most common Type III error: mixing up Type I and Type II!

Types of errors in hypothesis tests

- Type I error: rejecting H_0 when it is true.
- Type II error: not rejecting H_0 when it is false
- The probability of Type I and Type II error are inversely related

Table: Types of errors

	H_0 is false	H_0 is true
Reject H_0	Correct decision	Type I error (w.p. α)
Fail to reject H_0	Type II error	Correct decision

- Statistics joke from Casella and Berger: one ought not forget the most common Type III error: mixing up Type I and Type II!

Types of errors in hypothesis tests

- Type I error: rejecting H_0 when it is true.
- Type II error: not rejecting H_0 when it is false
- The probability of Type I and Type II error are inversely related

Table: Types of errors

	H_0 is false	H_0 is true
Reject H_0	Correct decision	Type I error (w.p. α)
Fail to reject H_0	Type II error	Correct decision

- Statistics joke from Casella and Berger: one ought not forget the most common Type III error: mixing up Type I and Type II!

Types of errors in hypothesis tests

- Type I error: rejecting H_0 when it is true.
- Type II error: not rejecting H_0 when it is false
- The probability of Type I and Type II error are inversely related

Table: Types of errors

	H_0 is false	H_0 is true
Reject H_0	Correct decision	Type I error (w.p. α)
Fail to reject H_0	Type II error	Correct decision

- Statistics joke from Casella and Berger: one ought not forget the most common Type III error: mixing up Type I and Type II!

Small-sample inference

- If we have small sample size, then the central limit theorem is not informative.
- Sampling distribution is no longer approximately normal but has a t-distribution.
- t-distribution has fatter tails than the standard normal the spread depending on degrees of freedom (df), where df$= n$–1. As n and df increase, standard deviation becomes more precise estimator of σ and t-distribution becomes less spread out and approaches normal.
- The two are practically identical when df ≥ 30
- t-distribution table lists t critical values corresponding to several one-tail probabilities and different df; as df increases.

Small-sample inference

- If we have small sample size, then the central limit theorem is not informative.
- Sampling distribution is no longer approximately normal but has a t-distribution.
- t-distribution has fatter tails than the standard normal the spread depending on degrees of freedom (df), where df$= n$–1. As n and df increase, standard deviation becomes more precise estimator of σ and t-distribution becomes less spread out and approaches normal.
- The two are practically identical when df ≥ 30
- t-distribution table lists t critical values corresponding to several one-tail probabilities and different df; as df increases.

Small-sample inference

- If we have small sample size, then the central limit theorem is not informative.
- Sampling distribution is no longer approximately normal but has a t-distribution.
- t-distribution has fatter tails than the standard normal the spread depending on degrees of freedom (df), where df$= n$–1. As n and df increase, standard deviation becomes more precise estimator of σ and t-distribution becomes less spread out and approaches normal.
- The two are practically identical when df $\geq$ 30
- t-distribution table lists t critical values corresponding to several one-tail probabilities and different df; as df increases.

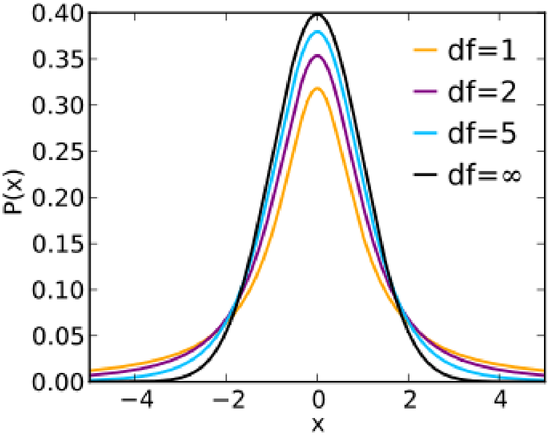
Small-sample inference

- If we have small sample size, then the central limit theorem is not informative.
- Sampling distribution is no longer approximately normal but has a t-distribution.
- t-distribution has fatter tails than the standard normal the spread depending on degrees of freedom (df), where df$= n$–1. As n and df increase, standard deviation becomes more precise estimator of σ and t-distribution becomes less spread out and approaches normal.
- The two are practically identical when df ≥ 30
- t-distribution table lists t critical values corresponding to several one-tail probabilities and different df; as df increases.

Small-sample inference

- If we have small sample size, then the central limit theorem is not informative.
- Sampling distribution is no longer approximately normal but has a t-distribution.
- t-distribution has fatter tails than the standard normal the spread depending on degrees of freedom (df), where df$= n–1$. As n and df increase, standard deviation becomes more precise estimator of σ and t-distribution becomes less spread out and approaches normal.
- The two are practically identical when df ≥ 30
- t-distribution table lists t critical values corresponding to several one-tail probabilities and different df; as df increases.

Small-sample inference



Small-sample inference

- The confidence intervals and test statistic for the mean resemble the large-sample size ones, but using the t-distribution and assuming that the variable of interest is distributed normally in the population.
- Confidence interval for the mean:

$$\mu = \overline{Y} \pm t_{\alpha/2} \times \frac{s}{\sqrt{n}}$$

  for a t-value defined by df$= n - 1$
- The test statistic is the t-statistic is this case:

$$t = \frac{\overline{Y} - \mu_0}{s/\sqrt{n}}$$

  where μ_0 is value hypothesised in H_0.
- The corresponding P-value, which can be one-tailed or two-tailed depending on H_A, derives from t-distribution table.

Small-sample inference

- The confidence intervals and test statistic for the mean resemble the large-sample size ones, but using the t-distribution and assuming that the variable of interest is distributed normally in the population.
- Confidence interval for the mean:

$$\mu = \overline{Y} \pm t_{\alpha/2} \times \frac{s}{\sqrt{n}}$$

  for a t-value defined by df$= n - 1$

- The test statistic is the t-statistic is this case:

$$t = \frac{\overline{Y} - \mu_0}{s/\sqrt{n}}$$

  where μ_0 is value hypothesised in H_0.

- The corresponding P-value, which can be one-tailed or two-tailed depending on H_A, derives from t-distribution table.

Small-sample inference

- The confidence intervals and test statistic for the mean resemble the large-sample size ones, but using the *t*-distribution and assuming that the variable of interest is distributed normally in the population.
- Confidence interval for the mean:

$$\mu = \overline{Y} \pm t_{\alpha/2} \times \frac{s}{\sqrt{n}}$$

  for a *t*-value defined by df$= n - 1$
- The test statistic is the *t*-statistic is this case:

$$t = \frac{\overline{Y} - \mu_0}{s/\sqrt{n}}$$

  where μ_0 is value hypothesised in H_0.
- The corresponding *P*-value, which can be one-tailed or two-tailed depending on H_A, derives from *t*-distribution table.

Small-sample inference

- The confidence intervals and test statistic for the mean resemble the large-sample size ones, but using the t-distribution and assuming that the variable of interest is distributed normally in the population.
- Confidence interval for the mean:

$$\mu = \overline{Y} \pm t_{\alpha/2} \times \frac{s}{\sqrt{n}}$$

  for a t-value defined by df$= n - 1$
- The test statistic is the t-statistic is this case:

$$t = \frac{\overline{Y} - \mu_0}{s/\sqrt{n}}$$

  where μ_0 is value hypothesised in H_0.
- The corresponding P-value, which can be one-tailed or two-tailed depending on H_A, derives from t-distribution table.

Self-study

- Hypothesis test for a proportion
  - Agresti and Finlay (1997) pp. 167-171.
  - The same principles, the same steps. Only the formula for the calculation of z-score (or t-statistic) is different.

Question (a) I

A researcher is assigned to work on a project investigating education of children in a Sub Saharan African country. She collected a random sample of 202 children from a national survey. The mean year of schooling of the sample was 4.5 with the standard deviation of 1.8.
(i) Is this significantly different from the mean years of schooling of all SSA countries at 5.2? Test the hypothesis at 1% significance level. [15%]

Answer

$\mu = 5.2$, $\bar{X} = 4.5$, $s = 1.8$
$n = 202 \Rightarrow$ large sample (>30) $\Rightarrow$ Normal Distribution
Test steps for full marks:

1. Assumptions
    1. Random sample
2. Hypothesis
    1. $H_0 : \mu = 5.2$
    2. $H_1 : \mu \neq 5.2$

Question (a) II

3. Test statistic
   $z = \frac{\bar{X}-\mu}{s/\sqrt{n-1}} = \frac{4.5-5.2}{1.8/\sqrt{202-1}} = -5.513$
4. Given $\alpha = 0.01$ (two-sided test)
   $z_{critical} = z_{0.005} = 2.58$
   Note: Use $\frac{s}{\sqrt{n}}$ is also acceptable for large samples. The two textbooks give 2 different formulas.
5. Decision
   $|z| > z_{critical}$
   $\therefore$ Reject H_0
6. Conclusions

Question (a) I

A researcher is assigned to work on a project investigating education of children in a Sub Saharan African country. She collected a random sample of 202 children from a national survey. The mean year of schooling of the sample was 4.5 with the standard deviation of 1.8.
(ii) Construct and interpret a 99% confidence interval for the mean years of schooling. [15%]

Answer

First, calculate the confidence interval:

$$\begin{aligned} CI = \bar{X} \pm z_{\alpha/2} * SE &= 4.5 \pm 2.58 * (1.8/\sqrt{202-1}) \\ &= 4.5 \pm 0.33 \\ &= [4.17,\ 4.83] \end{aligned}$$

Summary: interpretation of *CI* - the interval that contains the parameter (mean) with 99% confidence level.

Question (b) I

The researcher intends to further explore whether Sub Saharan Africa (SSA) is marginalised in the world, not only economically both also socio-politically. She calculated the mean of several socio-political indicators of several country groups based on UNDP data for the year 2001. The estimated results are reported in the table below.
(i) Based on the evidence from this table, can the researcher make a rigorous conclusion that SSA is marginalised economically and socio-politically? Explain.
Answer: Key Points

- Interpretation of the table
- Compare SSA with the rest, with reference to the research question. (SSA is marginalised economically, and socio-politically)
- Discuss whether evidence from this table is sufficient for a *rigorous* conclusion

Question (b) II

- Suggest appropriate statistical tests for a rigorous test. (At least one test is required)
  - A full mark requires a full discussion of the methods, including pros and cons.

Question (b) I

The researcher intends to further explore whether Sub Saharan Africa (SSA) is marginalised in the world, not only economically both also socio-politically. She calculated the mean of several socio-political indicators of several country groups based on UNDP data for the year 2001. The estimated results are reported in the table below.
(ii) Assume there are several countries in Arab States, South Asia and OECD which have extremely large values of GDP per capita. What are the consequences if the researcher uses mean GDP per capita of each region for his comparative analysis?

- Discuss the consequences of outliers for the mean
- How the bias may affect the comparative analysis.
- The use of median (we do not cover it in this course, but you can use the one sample median test). In Stata this is:

```
signrank life=48.8
```

Question (b) I

The researcher intends to further explore whether Sub Saharan Africa (SSA) is marginalised in the world, not only economically both also socio-politically. She calculated the mean of several socio-political indicators of several country groups based on UNDP data for the year 2001. The estimated results are reported in the table below.

(iii) The UNDP database has various economic and social indicators of all the countries in the world for the years of 1991 and 2001. Using this data, what statistical test(s) can the researcher employ to test the hypothesis that globalisation has played a role in the economic marginalisation of SSA? Give details of at least one method and discuss its limitations. You are not required to carry out the actual computation. [30%]

- You are required to show a sense of research design.
- The capability to turn a research question in to a **testable hypothesis**.
- Identify at least one appropriate method.

Question (b) II

- Operations of the test
  - The choice of indicators/variables to measure economic marginalisation
  - The measurement of globalisation (e.g. trade openness)
  - The test, the logic, how does it work?
- The limitations in method, data, measurement etc.

Note: The tests can be non-parametric or parametric e.g. t-test, ANOVA, regression. Good answers may use some example data, interpret the results in assumed example.