Readiness assessment questions

Q1- Rejecting the null hypothesis when $H_0$ is true is a

A. Type II error
B. Power
C. Correct decision
D. Type I error
Readiness assessment questions

Q1- Rejecting the null hypothesis when \( H_0 \) is true is a

The probability of rejecting the null hypothesis when \( H_0 \) is true is the type I error (alpha error, or false-positive error), therefore answer D is correct.

Q2- The ability of statistical test to detect a difference if it actually exists is

A. \( 1-\alpha \)
B. Type I error
C. \( 1-\beta \)
D. Type II error
Readiness assessment questions

Q2- The ability of statistical test to detect a difference if it actually exists is

The ability of statistical test to detect a difference if it actually exists is the power of the test. Power is the complement of beta (1-β), therefore answer C is correct.

---

Readiness assessment questions

Q3- If the average level of blood sodium levels among 150 medical students is 140 mEq/L, with a 95% CI 135-145 mEq/L. The 95% confidence interval represents:

A. The sample mean of 140 ± one standard deviation
B. We are 95% confident that the mean of other samples drawn from the population will fall within the given interval
C. The range of values for which there is only a 5% chance of not containing the true sample mean
D. If the interval is estimated for a single sample, we can be 95% confident that it includes the unknown population mean
Readiness assessment questions

Q4: If the average level of blood sodium levels among 36 diabetes insidious patients is 160 mEq/L. If the standard deviation for the population of all diabetes insidious patients is 12 mEq/L. The 95% confidence interval of the average population blood sodium levels is:

A. 124.00-196.00 mEq/L  
B. 140.00-180.00 mEq/L  
C. 158.04-161.96 mEq/L  
D. 156.08-163.92 mEq/L

$$100(1-\alpha) \text{ confidence interval of the mean}= \text{mean} \pm z_{(1-\alpha/2)} \times (SE)$$

$$95\% \text{CI of mean}= \text{mean} \pm z_{(0.975)} \times (SE) = \text{mean} \pm 1.96(\text{SE})$$

$$\text{SE of the mean} = \frac{SD}{\sqrt{N}}$$

$$95\% \text{ CI of the mean}= 160 \pm 1.96(\frac{12}{\sqrt{36}}) = 160 \pm 1.96(2)$$

$$95\% \text{ CI of the mean}= 160 \pm 3.92 = \textbf{156.08-163.92 mEq/L}$$
Readiness assessment questions

<table>
<thead>
<tr>
<th>Z</th>
<th>0.00</th>
<th>0.01</th>
<th>0.02</th>
<th>0.03</th>
<th>0.04</th>
<th>0.05</th>
<th>0.06</th>
<th>0.07</th>
<th>0.08</th>
<th>0.09</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>0.5000</td>
<td>0.5040</td>
<td>0.5080</td>
<td>0.5120</td>
<td>0.5160</td>
<td>0.5199</td>
<td>0.5239</td>
<td>0.5279</td>
<td>0.5319</td>
<td>0.5359</td>
</tr>
<tr>
<td>0.1</td>
<td>0.5398</td>
<td>0.5438</td>
<td>0.5478</td>
<td>0.5517</td>
<td>0.5557</td>
<td>0.5596</td>
<td>0.5636</td>
<td>0.5675</td>
<td>0.5714</td>
<td>0.5753</td>
</tr>
<tr>
<td>0.2</td>
<td>0.5793</td>
<td>0.5832</td>
<td>0.5871</td>
<td>0.5910</td>
<td>0.5948</td>
<td>0.5987</td>
<td>0.6026</td>
<td>0.6064</td>
<td>0.6103</td>
<td>0.6141</td>
</tr>
<tr>
<td>0.3</td>
<td>0.6179</td>
<td>0.6217</td>
<td>0.6255</td>
<td>0.6293</td>
<td>0.6331</td>
<td>0.6368</td>
<td>0.6406</td>
<td>0.6443</td>
<td>0.6480</td>
<td>0.6517</td>
</tr>
<tr>
<td>0.4</td>
<td>0.6554</td>
<td>0.6591</td>
<td>0.6628</td>
<td>0.6664</td>
<td>0.6700</td>
<td>0.6736</td>
<td>0.6772</td>
<td>0.6808</td>
<td>0.6844</td>
<td>0.6879</td>
</tr>
<tr>
<td>0.5</td>
<td>0.6915</td>
<td>0.6950</td>
<td>0.6985</td>
<td>0.7019</td>
<td>0.7054</td>
<td>0.7088</td>
<td>0.7123</td>
<td>0.7157</td>
<td>0.7190</td>
<td>0.7224</td>
</tr>
<tr>
<td>0.6</td>
<td>0.7257</td>
<td>0.7281</td>
<td>0.7304</td>
<td>0.7324</td>
<td>0.7347</td>
<td>0.7380</td>
<td>0.7422</td>
<td>0.7454</td>
<td>0.7486</td>
<td>0.7517</td>
</tr>
<tr>
<td>0.7</td>
<td>0.7580</td>
<td>0.7601</td>
<td>0.7615</td>
<td>0.7627</td>
<td>0.7638</td>
<td>0.7648</td>
<td>0.7657</td>
<td>0.7667</td>
<td>0.7675</td>
<td>0.7683</td>
</tr>
<tr>
<td>0.8</td>
<td>0.7881</td>
<td>0.7900</td>
<td>0.7919</td>
<td>0.7939</td>
<td>0.7956</td>
<td>0.7973</td>
<td>0.7989</td>
<td>0.8005</td>
<td>0.8019</td>
<td>0.8033</td>
</tr>
<tr>
<td>0.9</td>
<td>0.8159</td>
<td>0.8186</td>
<td>0.8212</td>
<td>0.8238</td>
<td>0.8264</td>
<td>0.8289</td>
<td>0.8315</td>
<td>0.8340</td>
<td>0.8365</td>
<td>0.8389</td>
</tr>
<tr>
<td>1.0</td>
<td>0.8413</td>
<td>0.8438</td>
<td>0.8461</td>
<td>0.8485</td>
<td>0.8508</td>
<td>0.8531</td>
<td>0.8554</td>
<td>0.8577</td>
<td>0.8599</td>
<td>0.8621</td>
</tr>
<tr>
<td>1.1</td>
<td>0.8643</td>
<td>0.8665</td>
<td>0.8686</td>
<td>0.8708</td>
<td>0.8729</td>
<td>0.8749</td>
<td>0.8770</td>
<td>0.8790</td>
<td>0.8810</td>
<td>0.8830</td>
</tr>
<tr>
<td>1.2</td>
<td>0.8849</td>
<td>0.8869</td>
<td>0.8888</td>
<td>0.8907</td>
<td>0.8925</td>
<td>0.8944</td>
<td>0.8962</td>
<td>0.8979</td>
<td>0.8997</td>
<td>0.9013</td>
</tr>
<tr>
<td>1.3</td>
<td>0.9032</td>
<td>0.9049</td>
<td>0.9066</td>
<td>0.9082</td>
<td>0.9099</td>
<td>0.9115</td>
<td>0.9131</td>
<td>0.9147</td>
<td>0.9162</td>
<td>0.9177</td>
</tr>
<tr>
<td>1.4</td>
<td>0.9092</td>
<td>0.9107</td>
<td>0.9122</td>
<td>0.9136</td>
<td>0.9149</td>
<td>0.9163</td>
<td>0.9176</td>
<td>0.9189</td>
<td>0.9200</td>
<td>0.9200</td>
</tr>
<tr>
<td>1.5</td>
<td>0.9232</td>
<td>0.9245</td>
<td>0.9257</td>
<td>0.9268</td>
<td>0.9279</td>
<td>0.9289</td>
<td>0.9300</td>
<td>0.9310</td>
<td>0.9319</td>
<td>0.9329</td>
</tr>
<tr>
<td>1.6</td>
<td>0.9422</td>
<td>0.9434</td>
<td>0.9446</td>
<td>0.9456</td>
<td>0.9466</td>
<td>0.9475</td>
<td>0.9484</td>
<td>0.9492</td>
<td>0.9500</td>
<td>0.9499</td>
</tr>
<tr>
<td>1.7</td>
<td>0.9554</td>
<td>0.9564</td>
<td>0.9573</td>
<td>0.9582</td>
<td>0.9591</td>
<td>0.9599</td>
<td>0.9608</td>
<td>0.9616</td>
<td>0.9625</td>
<td>0.9603</td>
</tr>
<tr>
<td>1.8</td>
<td>0.9641</td>
<td>0.9649</td>
<td>0.9656</td>
<td>0.9664</td>
<td>0.9671</td>
<td>0.9678</td>
<td>0.9686</td>
<td>0.9693</td>
<td>0.9699</td>
<td>0.9706</td>
</tr>
<tr>
<td>1.9</td>
<td>0.9734</td>
<td>0.9740</td>
<td>0.9746</td>
<td>0.9752</td>
<td>0.9758</td>
<td>0.9764</td>
<td>0.9769</td>
<td>0.9775</td>
<td>0.9781</td>
<td>0.9787</td>
</tr>
<tr>
<td>2.0</td>
<td>0.9772</td>
<td>0.9778</td>
<td>0.9783</td>
<td>0.9788</td>
<td>0.9793</td>
<td>0.9798</td>
<td>0.9803</td>
<td>0.9808</td>
<td>0.9812</td>
<td>0.9817</td>
</tr>
</tbody>
</table>

Type I error

- **Definition:**
  - Rejecting the null hypothesis when it is true
  - Accepting the alternative hypothesis when it is false
- **aka.**
  - Alpha error, False-positive error
- **Would like very small chance of rejecting the null when it is true** (to minimize harm)
Type II error

- Definition:
  - Accepting the null hypothesis when it is false
  - Rejecting the alternative hypothesis when it is true
- Aka.
  - Beta error
  - False-negative error
- Generally we would like small chance of rejecting the alternative when it is true, 20% is acceptable

<table>
<thead>
<tr>
<th>Decision of statistical test</th>
<th>Truth</th>
<th>Statistical test results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reject $H_0$</td>
<td>$H_o$ False (there is a difference)</td>
<td>Correct Power</td>
</tr>
<tr>
<td>Fail to reject $H_0$</td>
<td>$H_o$ True (there is no difference)</td>
<td>Type I error False-positive $\alpha=P(\text{reject } H_0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Decision of statistical test</th>
<th>Truth</th>
<th>Statistical test results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reject $H_0$</td>
<td>$H_o$ False (there is a difference)</td>
<td>Correct</td>
</tr>
<tr>
<td>Fail to reject $H_0$</td>
<td>$H_o$ True (there is no difference)</td>
<td>Type II error False-negative $\beta=P(\text{accept } H_0</td>
</tr>
</tbody>
</table>

Power

- Definition:
  - Is the probability of detecting a difference if it actually exists
  - Is the complement of type II error
  - Is the true positive rate (sensitivity)
  - Power $= P(\text{reject } H_0| H_o \text{ false}) = 1 - \beta$
Power

• Definition:
  • Is the probability of detecting a difference if it actually exists
  • Is the complement of type II error
  • Is the true positive rate (sensitivity)
  • Power= P(reject H₀ | H₀ false) = 1 - β
  • 80% is acceptable
    – Interpretation: There is 80% chance that the statistical test will detect the proposed difference, given that difference exists

Power

• Depends on the true magnitude of the association
  – Strong association is easier to detect than a weak association
• Depends on the variance of the measure of effect
  – The lesser the variance in the measure of effect, the less difficult to detect it
• Increase with
  – Increasing effect size (RR)
  – Cohort study/RCT: increasing frequency of the outcome in source population (up to about 0.5)
  – Case-control study: increasing frequency of exposure in source population (up to about 0.5)
  – Decreasing variance in the measure of effect
  – Increasing sample size
  – Increasing significance level
Alpha level

- Decided upon prior to data collection
- The highest acceptable risk of committing type I error
  - 5% is acceptable
    - Arbitrary level
    - \( \leq 5\% \rightarrow \text{reject the null} \)
    - \( > 5\% \rightarrow \text{fail to reject the null} \)
  - Interpretation of .05 alpha level
    - The investigator is willing to take not more than
      5% risk of erroneously rejecting the null hypothesis when it is in fact true

P value

- The probability of observing statistic as extreme or more extreme than the observed statistic given that the null hypothesis is true

- The p value is obtained from the test of significance
Confidence interval (CI)

- **Definition:**
  - Range of values that describe uncertainty about an estimate
  - A set of parameter values most compatible with the data

- Another method of estimating population values and indicating significance

Confidence interval (CI)

- **General formula**
  \[
  100(1-\alpha) \text{ CI} = \text{estimate} \pm (\text{confidence coefficient} \times \text{standard error of the estimate})
  \]

  - If the population SD is known:
    - \(100(1-\alpha)\) confidence interval of the mean = \(\text{mean} \pm z_{(1-\alpha/2)} \times \text{(SE)}\)
    - eg. 95%CI of mean = \(\text{mean} \pm z_{(0.975)} \times (\text{SE}) = \text{mean} \pm 1.96(\text{SE})\)

  - If the population SD is unknown:
    - \(100(1-\alpha)\) confidence interval of the mean = \(\text{mean} \pm t_{(1-\alpha/2)} \times (\text{SE})\)

  - Single population proportion:
    - \(100(1-\alpha)\) confidence interval of the proportion = \(\hat{p} \pm z_{(1-\alpha/2)} \times (\text{SE})\)
Standard error

- Is the standard deviation of the distribution of the means
- Is used in calculating CI
- Is used in tests of statistical significance
- SE is dependent on the size of the sample
  - Increasing the size of the sample $\rightarrow$ decreasing the SE

- Standard error = SE
  - SE of proportion $= \sqrt{\frac{p(1-p)}{N}}$
  - SE of the mean $= SD / \sqrt{N}$

Confidence interval (CI)

- Interpretation
  - For a series of samples, all of the same sample size $n$ obtained from a population, and $100(1-\alpha)\%$ CI estimating population parameter are constructed for each sample, then the relative frequency with which these intervals contain the true population parameter is $100(1-\alpha)\%$.
    - Ex. 95% CI: is a set of parameter values formed by a procedure, which if used repeatedly, will contain the true parameter 95% of the time (Statistical analysis of epidemiologic data, text for Prev 720, LS Magder)

  - For a single interval obtained from a single sample, a $100(1-\alpha)\%$ CI signifies that the investigator can be $100(1-\alpha)\%$ confident that this interval contains the unknown population parameter
    - Ex. 90% CI means that the investigator can be 90% confident that this interval contains the unknown population parameter
Confidence interval (CI)

- The width of the CI depends on the sample size
  - Increasing sample size $\rightarrow$ decreasing the width of the CI

- The width of the CI depends on the standard error
  - Increasing SE $\rightarrow$ increasing the width of the CI

- The width of the CI depends on the selected percentage of confidence
  - Increasing the confidence percentage $\rightarrow$ increasing the width of the CI e.g.. 95% CI is wider than a 90% CI

Application questions

Q1-3: A new antihypertensive drug was used in a clinical trial in 10 volunteers for 6 months. Baseline and 6-month post baseline systolic blood pressure are as follows:

<table>
<thead>
<tr>
<th>Subject</th>
<th>Systolic blood pressure (mm Hg)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline</td>
</tr>
<tr>
<td>1</td>
<td>160.00</td>
</tr>
<tr>
<td>2</td>
<td>190.00</td>
</tr>
<tr>
<td>3</td>
<td>185.00</td>
</tr>
<tr>
<td>4</td>
<td>160.00</td>
</tr>
<tr>
<td>5</td>
<td>175.00</td>
</tr>
<tr>
<td>6</td>
<td>180.00</td>
</tr>
<tr>
<td>7</td>
<td>165.00</td>
</tr>
<tr>
<td>8</td>
<td>173.00</td>
</tr>
<tr>
<td>9</td>
<td>168.00</td>
</tr>
<tr>
<td>10</td>
<td>155.00</td>
</tr>
</tbody>
</table>
Application questions

Q1: Without doing calculations, inspect the data, what can you conclude about the difference in baseline and 6-month post baseline systolic blood pressure values? (N.B. normal systolic blood pressure value is <120 mm Hg)

A. **Even if statistically significant, the difference is probably not clinically significant**

B. Even if clinically significant, the difference cannot be statistically significant

C. If statistically significant, the difference must be clinically significant

D. If clinically significant, the difference must be statistically significant

---

Answer A is correct because, inspecting the data suggests that the anti-hypertensive drug effect might not be clinically important, since blood pressure levels did not become normal. Answer B is incorrect because we cannot judge on statistical significance merely by the inspection of the data. Clinical significance and statistical significance are not the same and are not dependent, therefore answers C&D are incorrect. Also, a statistically significant difference could occur due to the large sample size although the effect size is small and have no clinical importance (C). In addition, statistical significance depends on the sample size, therefore, a clinically significant effect might not be statistically significant because of the lack of power (D).
Q2: If the investigator decided to use a one-tailed instead of two-tailed test of significance, how this would affect the study?

A. Does not affect statistical significance, but does affect power
B. Does not affect statistical significance, and does not affect power
C. **Does affect the statistical significance, and does affect the power**
D. Does affect statistical significance, but does not affect power

**Application questions**

Q2: The correct answer is C. The choice of a one-tailed or a two-tailed test affects the resultant p value and thus the significance and the power. This is because, in one-tailed test, the null hypothesis rejection area i.e. the extreme 5% (if alpha is set at 0.05) of the distribution is all to one side, therefore statistical significance (i.e., a p value less than alpha) is easier to attain and smaller than the two-tailed, which also increase the power. In contrast, in a two-tailed test of significance, the 5% rejection region (if alpha is set at 0.05) is divided into two halves of 0.025 each, at the two extreme ends of the curve therefore statistical significance is more difficult to attain thus the power of the test will decrease.
4- Assess the evidence

Two-tailed test
When the hypothesis is “difference in both directions”
A>B and A<B

One-tailed test
When the hypothesis is “difference in one direction”
A>B or B>A

Figure modified from a graph generated by http://www.imathas.com/stattools/norm.html

Application questions

Q3: Upon conducting formal statistical testing, the results indicated that there is a significant difference between the baseline and 6-month post baseline systolic blood pressure values such that the drug lowered systolic blood pressure by an average of 7 mm Hg (95% CI 4.2-9.8) (p-value of <0.0001). If the investigator would calculate 99% CI instead of the 95% CI, what changes do you expect?

A. The standard error will increase
B. **The width of the CI will increase**
C. The width of the CI will decrease
D. The average change in systolic blood pressure will decrease
Q3: Upon conducting formal statistical testing, the results indicated that there is a significant difference between the baseline and 6-month post baseline systolic blood pressure values such that the drug lowered systolic blood pressure by an average of 7 mm Hg (95% CI 4.2, 9.8) (p-value of <0.0001). If the investigator would calculate 99% CI instead of the 95% CI, what changes do you expect?

The width of the CI depends on the selected percentage of confidence. Increasing the confidence percentage \( \rightarrow \) increasing the width of the CI therefore 99% CI is wider than a 95% CI, because the confidence coefficient will be larger

\[
100(1-\alpha)\% \text{ CI} = \text{estimate} + (\text{confidence coefficient} \times \text{standard error of the estimate})
\]

Q4: The investigator decided to redo the study using a sample size of 100 patients. Assuming that the average change in systolic blood pressure and its standard deviation remained the same as in the study of 10 subjects, what changes would you expect?

A. The beta will increase
B. **The power will increase**
C. The standard error will increase
D. The width of the 95% CI will increase
Application questions

Q4: The investigator decided to redo the study using a sample size of 100 patients. Assuming that the average change in systolic blood pressure and its standard deviation remained the same as in the study of 10 subjects, what changes would you expect?

Given everything remained the same, increasing the sample size will lead to an increase in the study power (1-beta), therefore answer B is correct. Because beta is the complement of the power, then beta will decrease with the increased sample size thus A is incorrect. Increasing the sample size decreases the SE, therefore C is incorrect. The width of the 95% CI depends on the SE, which will decrease therefore the width of the 95% CI will decrease, therefore D is incorrect.