Applied Hypothesis Testing – Quiz

Probability & Statistics

(c) Joerg Osterrieder 2025

Question 1

What is a Type I error?

A. Failing to reject a false null hypothesis
B. Rejecting a true null hypothesis
C. Accepting the alternative hypothesis correctly
D. Making a calculation error

Question 1

What is a Type I error?

A. Failing to reject a false null hypothesis
B. Rejecting a true null hypothesis
C. Accepting the alternative hypothesis correctly
D. Making a calculation error

Answer: B
A Type I error (false alarm) occurs when we reject a true null hypothesis. This is controlled by the significance level alpha.

(c) Joerg Osterrieder 2025

Question 2

When should you use a paired t-test instead of a two-sample t-test?

A. When sample sizes are unequal
B. When comparing before-after measurements on the same subjects
C. When populations have different variances
D. When the sample size is large

Question 2

When should you use a paired t-test instead of a two-sample t-test?

A. When sample sizes are unequal
B. When comparing before-after measurements on the same subjects
C. When populations have different variances
D. When the sample size is large

Answer: B
Paired t-tests are used when measurements are taken on the same subjects (before-after), matched pairs, or repeated measures. This design controls for individual differences.

Question 3

What does the F-statistic in ANOVA represent?

A. The ratio of within-group to between-group variance
B. The ratio of between-group to within-group variance
C. The total variance in the data
D. The sample size divided by the number of groups

Question 3

What does the F-statistic in ANOVA represent?

A. The ratio of within-group to between-group variance
B. The ratio of between-group to within-group variance
C. The total variance in the data
D. The sample size divided by the number of groups

Answer: B
The F-statistic is the ratio of between-group variance (signal) to within-group variance (noise). A large F indicates groups differ more than expected by chance.

Question 4

Why is the Bonferroni correction used?

A. To increase statistical power
B. To control the family-wise error rate with multiple comparisons
C. To calculate effect sizes
D. To test normality assumptions

Question 4

Why is the Bonferroni correction used?

A. To increase statistical power
B. To control the family-wise error rate with multiple comparisons
C. To calculate effect sizes
D. To test normality assumptions

Answer: B
The Bonferroni correction divides alpha by the number of tests to control the family-wise error rate. Without it, multiple tests inflate the chance of a false positive.

(c) Joerg Osterrieder 2025

Question 5

Statistical power is defined as:

A. The probability of making a Type I error
B. The probability of correctly rejecting a false null hypothesis
C. The significance level of the test
D. The effect size divided by the sample size

Question 5

Statistical power is defined as:

A. The probability of making a Type I error
B. The probability of correctly rejecting a false null hypothesis
C. The significance level of the test
D. The effect size divided by the sample size

Answer: B
Power = 1 - beta is the probability of detecting a real effect (correctly rejecting a false H0). We aim for power of at least 0.80.

Question 6

Cohen's d = 0.5 indicates:

A. A small effect size
B. A medium effect size
C. A large effect size
D. No effect

Question 6

Cohen's d = 0.5 indicates:

A. A small effect size

B. A medium effect size

C. A large effect size

D. No effect

Answer: B

Cohen's guidelines: d = 0.2 is small, d = 0.5 is medium, d = 0.8 is large. Cohen's d measures the difference between means in standard deviation units.

Question 7

Tukey HSD test is used to:

A. Test normality assumptions
B. Compare all pairwise group means after ANOVA
C. Calculate the F-statistic
D. Estimate sample size for power analysis

Question 7

Tukey HSD test is used to:

A. Test normality assumptions
B. Compare all pairwise group means after ANOVA
C. Calculate the F-statistic
D. Estimate sample size for power analysis

Answer: B
Tukey's Honestly Significant Difference test performs all pairwise comparisons between group means while controlling the family-wise error rate.

(c) Joerg Osterrieder 2025

Question 8

Which of the following increases statistical power?

A. Decreasing sample size
B. Decreasing effect size
C. Increasing sample size
D. Using a smaller significance level

(c) Joerg Osterrieder 2025

Question 8

Which of the following increases statistical power?

A. Decreasing sample size
B. Decreasing effect size
C. Increasing sample size
D. Using a smaller significance level

Answer: C
Larger sample sizes increase power by reducing sampling variability. Larger effect sizes and higher alpha also increase power.

(c) Joerg Osterrieder 2025

Question 9

One-way ANOVA tests:

A. Whether the means of two groups are equal
B. Whether at least one group mean differs from the others
C. Which specific groups differ from each other
D. Whether variances are equal across groups

Question 9

One-way ANOVA tests:

A. Whether the means of two groups are equal
B. Whether at least one group mean differs from the others
C. Which specific groups differ from each other
D. Whether variances are equal across groups

Answer: B
ANOVA tests whether at least one group mean differs from the others. It does not tell which groups differ - that requires post-hoc tests.

(c) Joerg Osterrieder 2025

Question 10

What is the relationship between p-value and effect size?

A. A small p-value always means a large effect size
B. P-value and effect size measure the same thing
C. A small p-value can occur with a small effect if sample size is large
D. Effect size is always larger than p-value

Question 10

What is the relationship between p-value and effect size?

A. A small p-value always means a large effect size

B. P-value and effect size measure the same thing

C. A small p-value can occur with a small effect if sample size is large

D. Effect size is always larger than p-value

Answer: C

P-value measures evidence against H0; effect size measures magnitude. With large samples, even tiny effects can be statistically significant but practically unimportant.