Statistical Facts in Human Physiology and Health

Statistics · Answer Key · 16 Questions

1. In a large-scale epidemiological study investigating the incidence of a specific type of cancer, what is the most appropriate statistical measure to describe the rate at which new cases occur in a defined population over a specific time period?

A) Prevalence
B) Sensitivity
C) Incidence Rate
D) Positive Predictive Value

2. When analyzing the results of a clinical trial for a new medication designed to lower blood pressure, a statistically significant p-value (e.g., $p < 0.05$) typically indicates:

A) The treatment is definitely effective.
B) There is strong evidence against the null hypothesis.
C) The observed effect is solely due to chance.
D) The sample size was too small.

3. A meta-analysis combining data from multiple randomized controlled trials on a new vaccine's efficacy found a pooled odds ratio of 0.30 (95% CI: 0.20-0.45). What does this statistically significant result suggest about the vaccine?

A) The vaccine increases the risk of the disease.
B) The vaccine is associated with a reduced risk of the disease.
C) There is no discernible effect of the vaccine.
D) The studies included were too heterogeneous to draw conclusions.

4. In medical diagnostics, the area under the Receiver Operating Characteristic (ROC) curve (AUC) is a measure of a diagnostic test's ability to discriminate between:

A) Two groups of healthy individuals.
B) Two groups of individuals with the same disease.
C) Individuals with and without a specific condition.
D) Different stages of the same disease.

5. Survival analysis, commonly used in oncology, estimates the probability of an event (like death) occurring over time. The Kaplan-Meier estimator is a non-parametric statistic used to estimate:

A) The incidence of new cases.
B) The prevalence of existing cases.
C) The survival function.
D) The hazard ratio.

6. When assessing the reliability of a diagnostic test, a high specificity indicates that the test is good at:

A) Correctly identifying individuals with the disease.
B) Correctly identifying individuals without the disease.
C) Minimizing false positive results.
D) Minimizing false negative results.

7. In a study examining the relationship between exercise frequency and cardiovascular health, a Pearson correlation coefficient of -0.75 was calculated. This indicates:

A) A weak positive linear relationship.
B) A strong negative linear relationship.
C) No linear relationship.
D) A weak negative linear relationship.

8. The standard deviation is a statistical measure that quantifies:

A) The average of a dataset.
B) The central tendency of a dataset.
C) The spread or dispersion of data points around the mean.
D) The difference between the maximum and minimum values.

9. In pharmacokinetics, the concept of bioavailability, often expressed as a percentage, refers to the fraction of an administered dose of an unchanged drug that reaches:

A) The liver.
B) The stomach.
C) The systemic circulation.
D) The site of action.

10. When comparing the average blood glucose levels of two groups of patients (diabetic vs. non-diabetic), an independent samples t-test is a common statistical tool used to determine:

A) If the variances of the two groups are equal.
B) If there is a statistically significant difference between the means of the two groups.
C) The correlation between glucose levels and age.
D) The proportion of patients in each group.

11. The Poisson distribution is a discrete probability distribution that is often used to model the number of events occurring within a fixed interval of time or space. In a health context, it could be used to model:

A) The height of individuals in a population.
B) The number of births in a hospital ward per day.
C) The distribution of blood pressure values.
D) The weight of newborns.

12. In genetic epidemiology, linkage disequilibrium (LD) refers to the non-random association of alleles at different loci. High LD in a population suggests:

A) Alleles are inherited independently.
B) Alleles are often inherited together due to proximity on a chromosome or selective forces.
C) There are no genetic mutations.
D) The population is in Hardy-Weinberg equilibrium.

13. When a confidence interval for the mean difference between two groups does not include zero, it generally implies:

A) There is no statistically significant difference between the group means.
B) There is a statistically significant difference between the group means.
C) The sample size was insufficient.
D) The data are not normally distributed.

14. The concept of 'effect size' in statistical analysis is important because it quantifies:

A) The probability of rejecting a false null hypothesis.
B) The magnitude of the difference or relationship observed, independent of sample size.
C) The variability of the data within a group.
D) The proportion of variance explained by a statistical model.

15. In studies of chronic diseases, the 'hazard function' (or hazard rate) in survival analysis represents:

A) The probability of surviving up to a certain time point.
B) The cumulative probability of an event occurring by a specific time.
C) The instantaneous rate of an event occurring at a specific time, given survival up to that time.
D) The proportion of the population that has experienced the event.

16. When a statistical model is used to predict the risk of developing a disease based on multiple predictor variables (e.g., age, BMI, smoking status), the coefficients (often denoted as beta coefficients) in logistic regression represent:

A) The absolute change in risk for a one-unit increase in the predictor.

B) The change in the log-odds of the outcome for a one-unit increase in the predictor.

C) The probability of the outcome occurring.

D) The correlation between the predictor and the outcome.