# Discovering Statistics Using IBM SPSS Statistics, 5th Ed. Practice Test (Sample)

## Study Guide

BY EXAMZIFY

**Everything you need from our exam experts!**

# Table of Contents

# Introduction

**Preparing for a certification exam can feel overwhelming, but with the right tools, it becomes an opportunity to build confidence, sharpen your skills, and move one step closer to your goals. At Examzify, we believe that effective exam preparation isn't just about memorization, it's about understanding the material, identifying knowledge gaps, and building the test-taking strategies that lead to success.**

**This guide was designed to help you do exactly that.**

**Whether you're preparing for a licensing exam, professional certification, or entry-level qualification, this book offers structured practice to reinforce key concepts. You'll find a wide range of multiple-choice questions, each followed by clear explanations to help you understand not just the right answer, but why it's correct.**

**The content in this guide is based on real-world exam objectives and aligned with the types of questions and topics commonly found on official tests. It's ideal for learners who want to:**

- **Practice answering questions under realistic conditions,**
- **Improve accuracy and speed,**
- **Review explanations to strengthen weak areas, and**
- **Approach the exam with greater confidence.**

**We recommend using this book not as a stand-alone study tool, but alongside other resources like flashcards, textbooks, or hands-on training. For best results, we recommend working through each question, reflecting on the explanation provided, and revisiting the topics that challenge you most.**

**Remember: successful test preparation isn't about getting every question right the first time, it's about learning from your mistakes and improving over time. Stay focused, trust the process, and know that every page you turn brings you closer to success.**

**Let's begin.**

# How to Use This Guide

**This guide is designed to help you study more effectively and approach your exam with confidence. Whether you're reviewing for the first time or doing a final refresh, here’s how to get the most out of your Examzify study guide:**

### 1. Start with a Diagnostic Review

**Skim through the questions to get a sense of what you know and what you need to focus on. Your goal is to identify knowledge gaps early.**

### 2. Study in Short, Focused Sessions

**Break your study time into manageable blocks (e.g. 30 – 45 minutes). Review a handful of questions, reflect on the explanations.**

### 3. Learn from the Explanations

**After answering a question, always read the explanation, even if you got it right. It reinforces key points, corrects misunderstandings, and teaches subtle distinctions between similar answers.**

### 4. Track Your Progress

**Use bookmarks or notes (if reading digitally) to mark difficult questions. Revisit these regularly and track improvements over time.**

### 5. Simulate the Real Exam

**Once you're comfortable, try taking a full set of questions without pausing. Set a timer and simulate test-day conditions to build confidence and time management skills.**

### 6. Repeat and Review

**Don’t just study once, repetition builds retention. Re-attempt questions after a few days and revisit explanations to reinforce learning. Pair this guide with other Examzify tools like flashcards, and digital practice tests to strengthen your preparation across formats.**

**There’s no single right way to study, but consistent, thoughtful effort always wins. Use this guide flexibly, adapt the tips above to fit your pace and learning style. You've got this!**

# Questions

1. **Which statistical test uses the t-statistic to determine whether the means of two independent samples differ significantly?**

   **A. Independent t-test**

   **B. Paired t-test**

   **C. ANOVA**

   **D. Chi-square test**

2. **What is the standardized measure of the strength of the relationship between two variables when one variable is dichotomous?**

   **A. Biserial correlation**

   **B. Point-biserial correlation**

   **C. Pearson correlation**

   **D. Spearman correlation**

3. **Which statement describes grand mean centring?**

   **A. Subtract the grand mean of the variable from every score.**

   **B. Subtract the mean of the scores within the sample from the scores.**

   **C. Multiply each score by the grand mean.**

   **D. Subtract the minimum score from each score.**

4. **Hat values are another name for which regression diagnostic concept?**

   **A. Leverage**

   **B. Influential observations**

   **C. Standardized residuals**

   **D. Confidence intervals**

5. **Which statistic is most affected by extreme values?**

   **A. Mode**

   **B. Median**

   **C. Mean**

   **D. Range**

**6. Which statement about z-score transformation is true?**

- **A. It changes the shape of the distribution.**
- **B. It leaves the units unchanged.**
- **C. It does not standardize data.**
- **D. It standardizes data so that the transformed distribution has a mean of 0 and a standard deviation of 1.**

**7. Counterbalancing, in the simplest case of two conditions, is used to remove bias caused by what?**

- **A. Practice effects**
- **B. Placebo effects**
- **C. Demand characteristics**
- **D. Measurement error**

**8. In regression analyses, what is a predictor variable?**

- **A. The outcome variable.**
- **B. The residual.**
- **C. The coefficient.**
- **D. A variable that is used to predict values of the outcome variable.**

**9. Pairwise comparisons are used for which of the following?**

- **A. Comparisons of pairs of means.**
- **B. Comparisons among all group means simultaneously.**
- **C. Assessing relationships between two variables.**
- **D. Testing the equality of variances across groups.**

**10. Welch's F is used to accommodate what condition in comparing group means?**

- **A. Used when the assumption of equal variances across groups is violated.**
- **B. Used when sample sizes across groups are equal.**
- **C. Used when data are normal and variances are equal.**
- **D. Used to accommodate heterogeneity of variance, i.e., unequal variances across groups.**

# Answers

**1. A**
**2. A**
**3. A**
**4. A**
**5. C**
**6. D**
**7. A**
**8. D**
**9. A**
**10. D**

# Explanations

## 1. Which statistical test uses the t-statistic to determine whether the means of two independent samples differ significantly?

**A. Independent t-test**

**B. Paired t-test**

**C. ANOVA**

**D. Chi-square test**

**Testing whether two independent groups have different means is done with the independent samples t-test. It relies on a t-statistic that compares the observed difference between the group means to the variability within the groups. If the difference is large relative to this variability, the t-statistic is large in magnitude and the p-value is small, leading to rejection of the null hypothesis that the means are equal. This situation applies when you have two separate groups and a continuous outcome. The other tests fit different situations: a paired t-test is for related or matched samples, where you compare differences within pairs rather than two independent means; ANOVA is used when comparing means across three or more groups (the two-group case is mathematically a special case of ANOVA but is typically reported as a t-test); and the chi-square test handles categorical data, not means. If variances between the two groups are assumed equal, you use the pooled version; if not, Welch's t-test is used. In practice, SPSS will often report Levene's test to guide which version to use.**

## 2. What is the standardized measure of the strength of the relationship between two variables when one variable is dichotomous?

**A. Biserial correlation**

**B. Point-biserial correlation**

**C. Pearson correlation**

**D. Spearman correlation**

**When one variable is dichotomous (two categories) and the other is continuous, the standard way to quantify the strength of their relationship is the point-biserial correlation. This is a special case of Pearson correlation where the binary variable is coded as 0 and 1. It reflects how much the continuous variable's values differ between the two groups, standardized by the overall variability. A positive value means higher scores tend to occur in the group coded as 1, a negative value means higher scores tend to occur in the group coded as 0, and larger absolute values indicate a stronger association. Biserial correlation, by contrast, is used when the dichotomous variable is believed to arise from an underlying continuous trait that has been split into two categories. In everyday data with an observed 0/1 dichotomy, the point-biserial approach is the appropriate and widely used measure.**

## 3. Which statement describes grand mean centring?

**A. Subtract the grand mean of the variable from every score.**

**B. Subtract the mean of the scores within the sample from the scores.**

**C. Multiply each score by the grand mean.**

**D. Subtract the minimum score from each score.**

**Grand mean centering means adjusting each score by removing the overall average across all observations, so the transformed values are centered around zero. Specifically, you subtract the grand mean from every score, which makes the new variable have a mean of zero. This helps in regression, especially with interaction terms, because it makes the interpretation of main effects clearer and can reduce multicollinearity between predictors. The statement that describes subtracting the grand mean from every score is the correct description of this process. The other options describe different actions—centering by a within-sample mean (group-mean centering), multiplying by a constant, or subtracting the minimum value—which do not implement grand-mean centering.**

## 4. Hat values are another name for which regression diagnostic concept?

**A. Leverage**

**B. Influential observations**

**C. Standardized residuals**

**D. Confidence intervals**

**Leverage is the regression diagnostic concept represented by hat values. Hat values are the diagonal elements of the hat matrix H = X(X'X)^{-1}X', and they quantify how much an observation's predictor values influence its fitted value ŷ. Each h_ii shows how strongly observation i can affect the regression line: observations with large hat values have predictor values that are far from the center of the data, giving them more potential to sway the estimates. High leverage is about the potential to influence, and it doesn't by itself guarantee influence unless accompanied by a large residual. Hat values typically range from 0 to 1, with larger values signaling greater leverage. In practice, recognizing high-leverage points helps assess their possible impact on the model's fit and coefficients. Standardized residuals relate to residual size, not how far the X-values lie from the typical pattern; influential observations depend on both leverage and residuals; confidence intervals pertain to precision of estimates rather than the influence of individual points.**

## 5. Which statistic is most affected by extreme values?

**A. Mode**

**B. Median**

**C. Mean**

**D. Range**

**The mean is the statistic most affected by extreme values because it uses every data point in its calculation. It is the balance point of the data, so an outlier pulls the average toward itself. For example, with data like 1, 2, 2, 3, 100, the mean becomes much larger than it would be without the outlier, while the median stays the same or changes only slightly. The mode depends on the most frequent value and often remains unchanged by an extreme value. The range changes directly because it stretches from the smallest to the largest value, so it can be influenced by extremes too, but the mean uniquely reflects changes across all observations rather than just the extremes.**

## 6. Which statement about z-score transformation is true?

**A. It changes the shape of the distribution.**

**B. It leaves the units unchanged.**

**C. It does not standardize data.**

**D. It standardizes data so that the transformed distribution has a mean of 0 and a standard deviation of 1.**

**Z-score transformation standardizes data by expressing each value as how many standard deviations it is from the mean. For each observation, you subtract the mean and divide by the standard deviation. This re-scaling makes the transformed distribution have a mean of zero and a standard deviation of one, while the shape of the distribution stays the same because you're only shifting and rescaling, not altering skewness or tails. As a result, the units disappear and the scores become dimensionless. If the original data are normally distributed, the z-scores follow a standard normal distribution with mean 0 and SD 1. This method is especially useful for comparing values across different scales or identifying outliers.**

## 7. Counterbalancing, in the simplest case of two conditions, is used to remove bias caused by what?

**A. Practice effects**

**B. Placebo effects**

**C. Demand characteristics**

**D. Measurement error**

**Counterbalancing in a two-condition design targets order effects—the bias introduced by the sequence in which conditions are presented. The simplest form is to have half the participants experience condition A first, then condition B, and the other half experience B first, then A. This spreads any learning or fatigue that comes from doing one task before the other across both conditions. As a result, improvements (practice effects) or declines due to fatigue balance out, so the comparison between the two conditions reflects the actual manipulation rather than the order in which they were encountered. Other biases like placebo effects, demand characteristics, or measurement error aren't what counterbalancing primarily fixes, so the most appropriate choice is practice effects.**

## 8. In regression analyses, what is a predictor variable?

**A. The outcome variable.**

**B. The residual.**

**C. The coefficient.**

**D. A variable that is used to predict values of the outcome variable.**

**In regression, a predictor variable is the variable used to predict values of the outcome variable. It serves as the input that helps explain or estimate what you're trying to forecast. The outcome (or dependent) variable is what you're predicting, so it's not the predictor itself. The residual represents the part of the outcome that the model doesn't explain—the difference between observed and predicted values. The coefficient is the estimated effect size that tells you how much the outcome changes with a one-unit change in the predictor, not the predictor itself. Predictor variables can be continuous or encoded categories (dummy variables) and are the inputs that drive the prediction.**

## 9. Pairwise comparisons are used for which of the following?

**A. Comparisons of pairs of means.**

**B. Comparisons among all group means simultaneously.**

**C. Assessing relationships between two variables.**

**D. Testing the equality of variances across groups.**

**Pairwise comparisons focus on differences between two means at a time. After an ANOVA shows that at least one group mean differs, pairwise tests examine each possible pair of groups (for example, group 1 vs group 2, group 1 vs group 3, and so on) to identify exactly which means differ. Because several comparisons are made, researchers usually apply adjustments (like Tukey or Bonferroni) to keep the overall chance of a false positive under control. This approach directly captures differences between two means at a time, which is what pairwise comparisons are designed to assess. The other ideas describe different analyses: comparing all group means at once is an omnibus or global test, assessing relationships between two variables refers to correlation or regression, and testing equality of variances concerns variances rather than means.**

**10. Welch's F is used to accommodate what condition in comparing group means?**

**A. Used when the assumption of equal variances across groups is violated.**

**B. Used when sample sizes across groups are equal.**

**C. Used when data are normal and variances are equal.**

**D. Used to accommodate heterogeneity of variance, i.e., unequal variances across groups.**

**Welch's F is used when the spread of scores differs across groups—that is, when variances are not equal from group to group (heterogeneity of variance). In standard ANOVA we assume equal variances, and when that assumption is violated, the regular F test can mislead. Welch's approach adjusts the degrees of freedom (via the Satterthwaite approximation) to provide a more accurate test of whether the group means differ under unequal variances and often unequal sample sizes. This makes it the appropriate choice specifically for handling variance heterogeneity across groups. The other scenarios—equal variances, equal sample sizes, or data that are normal with equal variances—don't require Welch's correction.**

# Next Steps

**Congratulations on reaching the final section of this guide. You've taken a meaningful step toward passing your certification exam and advancing your career.**

**As you continue preparing, remember that consistent practice, review, and self-reflection are key to success. Make time to revisit difficult topics, simulate exam conditions, and track your progress along the way.**

**If you need help, have suggestions, or want to share feedback, we'd love to hear from you. Reach out to our team at hello@examzify.com.**

**Or visit your dedicated course page for more study tools and resources:**

**https://discoveringstatsusingibmspss5e.examzify.com**

**We wish you the very best on your exam journey. You've got this!**