Part I True-False Statements: Fill in **T** or **F** (4 points each)

1. In a regression involving age and serum cholesterol level, the “least squares” or “best fit” line calculated from the sample data estimates the age-adjusted mean serum cholesterol level in the population. \_\_\_
2. If in a random sample of data we find the linear correlation coefficient, r, to be highly significant (p < 0.001), we may conclude that there is a dependent relationship between Y and X in the population. \_\_\_
3. The following data are from a case-control study investigating a possible relationship between radiation exposure and myeloid leukemia:

|  |  |  |
| --- | --- | --- |
|  | Case | Control |
| Exposed | 11 | 8 |
| Unexposed | 4 | 15 |

The appropriate procedure for evaluating the significance of such data is the Fisher exact test. \_\_\_

1. The maximum likelihood method is used to estimate parameters in logistic regression? \_\_\_
2. In a logistic regression model with 2 predictor variables (risk factors A and B) and no interaction term, the estimated relative risk (RR) for subjects exposed to both risk factor A (adjusted RR estimate = 2.0) and risk factor B (adjusted RR estimate = 1.5) is 3.5. \_\_\_

Part II Brief Responses (4 points each)

1. Part 1: For a fixed sample size: as the number of independent variables in a regression model increases, the power of the regression \_\_\_
2. increases
3. decreases

Part 2: When a potential confounding variable is found to affect the β-coefficient estimate for the variable of interest in a regression model and is therefore added to the model, the coefficient of determination (R2) \_\_\_

a) increases

1. decreases
2. An F-test for \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ must be performed prior to conducting an analysis of covariance (ANCOVA).

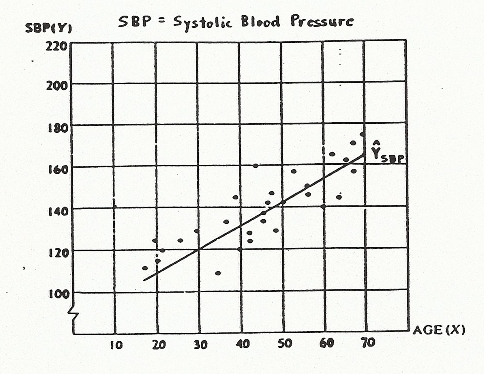
1. Part 1: Under what circumstances are chi-square tests biased? \_\_\_
2. if any expected value is less than 1.0 or > 25% of the expected values are less than 5.0
3. small sample size
4. when there is 1 degree of freedom
5. all of the above

Part 2: What is the best statistical test to apply when dealing with simple follow-up data (single 2 x 2 table)? \_\_\_

1. chi-square test
2. fisher exact test
3. binomial test
4. Refer to the figure below:

Part 1: What is the estimated mean SBP for age 30? \_\_\_

Part 2: What is the estimated mean SBP for age 80? \_\_\_



1. As health commissioner of city B you have decided to implement a $10 million exercise program if results from a recently concluded cohort study in 2,000 males aged 40-59 years showed there was 95% confidence (exact binomial) that the relative risk (RR) among subjects with a baseline resting heart rate (HR) > 80 beats/min vs. those with a baseline resting HR ≤ 80 beats/min was different from 1.50. The RR estimate was 1.55 with a 95% confidence interval of 1.31 to 1.91. What is your decision? \_\_\_

1. Implement the program.
2. Do not implement the program.
3. Part 1: Name one limitation when using the Mantel-Haenszel method to control for confounding in case-control studies.

Part 2: Name one limitation when using matching to control for confounding in a case-control study?

1. Name 2 types of censoring that can occur during follow-up in a cohort study or clinical trial.
2. A potential confounding variable should be included in a multiple proportional hazards regression model if \_\_\_
3. it is an independent risk factor for the outcome
4. it is a statistically significant predictor for the outcome
5. it differs in frequency between exposed and unexposed study subjects
6. choices a and c
7. choices a and b
8. all of the above
9. Using a box or arrow diagram, show how the total sum of squares is partitioned in a 1-way repeated measures ANOVA?

10. Match each description to the correct confidence interval in the diagram below. Must get all matches correct. No partial credit.

**Fail to reject H0 P> 0.05.** Not statistically significant despite large sample size, which justifies acceptance of the null hypothesis (CI captures 1 because the RR estimate is close to 1). \_\_\_

**Fail to reject H0 P> 0.05.** Although the RR estimate is close to 1, we cannot accept the null hypothesis since the CI is wide (imprecise). Cannot distinguish benefit from harm. \_\_\_

**Reject H0 and accept HA. P< 0.05.** Not clinically significant. RR estimate is statistically significant but the CI is wide (imprecise). The risk reduction may be large and real, but the upper limit of the CI is too close to 1 to take action. \_\_\_

**Fail to reject H0 P> 0.05.** Not statistically significant due to small sample size. This signifies a probable type II error (CI captures 1 not because the RR estimate is close to 1 but because the interval is wide). \_\_\_

**Reject H0 and accept HA. P< 0.05.** Clinically significant, large sample estimate, upper confidence limit far from null hypothesis RR, a blockbuster drug. \_\_\_

RR = 1

E

D

C

B

A

Benefit Harm

Part III. Problems (20 points each)

1. A study was performed to evaluate a potential relationship between heart rate and exercise minutes per week in a random sample of 11 subjects aged 20 to 29 years.

The results were as follows:

Exercise (min/week) HR (beats/min)

10 88

20 83

30 78

40 75

50 78

60 75

70 70

80 66

90 63

100 65

110 60

Using the attached SAS file, perform a simple linear regression analysis with the dependent variable, HR, and independent predictor variable, Exmin/week.

Assume normality and equal variance. Use α = 0.05 (two-tailed) and assume 80% power.

1. State the null and alternative hypotheses for testing the slope.
2. List the critical value for testing the slope.
3. List the R2.
4. Report your decision based on the critical value (reject Ho and accept Ha OR fail to reject Ho), P-value, and 95% confidence interval.
5. If the decision was to fail to reject Ho, can Ho be accepted?

2. In an epidemiologic study, 147 incident bladder cancer cases were compared to 179 controls selected from the population that gave rise to the cases with respect to smoking history. The data were as follows:

|  |  |  |  |
| --- | --- | --- | --- |
| Exposure status | Disease Status | |  |
| Case | Control | Total |
| Smoker | 90 | 82 | 172 |
| Nonsmoker | 57 | 97 | 154 |
| Total | 147 | 179 | 326 |

Perform the appropriate statistical test of the relationship between bladder cancer and smoking history using the attached SAS file. Assume α = 0.05 (two-tailed) and 80% power.

1. State the null and alternative hypotheses.
2. List the critical value.
3. List the estimated odds ratio.
4. Using the chi-square, P-value, and odds ratio output from SAS, report your decision based on the critical value (reject Ho and accept HA OR fail to reject Ho), P-value, and 95% confidence interval
5. If the decision was to fail to reject Ho, can Ho be accepted?