Solutions to Biostatistics and Epidemiology Step 1 Sample Questions Set 2

1. The correct answer is **A**. Recall that sensitivity is the probability of a positive test result among those that have the disease. Screening tests are generally designed to have high sensitivity in order to 'rule in' patients in order to not miss potential cases (see page 16 of Unit 5). Additionally, the consequences of failing to detect a case are severe, and the disease is easily treated once detected. For all these reasons, a highly sensitive test is desirable.
2. The correct answer is **A**. Notice that the question only pertains to people with the disease. Since the disease prevalence is 10%, then 10% of 1000 people = 100 people have the disease. Recall that sensitivity is the probability of a positive test result among those that have the disease. Therefore a sensitivity of 80% implies that 80 of the 100 cases will test positive, and the other 20 will incorrectly test negative.
3. The correct answer is **B**. The 'spread ' (or dispersion) associated with the distribution of the marker measurements based on the new test is smaller than that associated with the old test. Therefore, there will be fewer healthy subjects misclassified as diseased and fewer diseased subjects misclassified as healthy when the new test is used. This translates into higher sensitivity and higher specificity.
4. The correct answer is **B**. By shifting the test threshold further to the right, the number of healthy patients misclassified as diseased will go down, thereby decreasing the number of false positives and increasing the specificity.
5. The correct answer is **C**. A false negative is a diseased subject misclassified as healthy. In this problem, the prevalence is 1% so there are 1,000 patients with TB. Of these, 90% will be correctly classified and 10% will be misclassified. Therefore the number of false negatives is 100.
6. The correct answer is **D**. The first sentence in this problem tells us that the proportion of test positives among the diseased patients is 9 times higher than the proportion of test positives among the non-diseased patients. The proportion of test positives among the diseased patients is the definition of sensitivity. The proportion of test negatives among the non-diseased patients is specificity, and the proportion of test positives among the non-diseased patients is 1 – specificity. Therefore, the ratio described is sensitivity/(1 – specificity).
7. The correct answer is **D**. Since this patient tested negative, we are interested in evaluating the probability she doesn't have the disease given that she tested negative. This is the definition of negative predictive value. If NPV is high, then the probability of having the disease given a negative test result will be low.
8. The correct answer is **E**. The question is asking you to calculate 1 - negative predictive value – that is, the probability of having disease given that the test is negative. Based on the table, 90 subjects tested negative, and among these, 12 actually have the disease. So 1 – NPV= 12/90.
9. The correct answer is **B**. Length-time bias is a form of bias in screening tests that ascertain slowly progressing forms of the disease. The apparent increase in survival based on such a test is not due to the test actually improving survival but due to the fact that less aggressive forms of the disease are more readily diagnosed by the test.
10. The correct answer is **C**. The definition of a p-value is the probability of observing by chance alone a result as or more extreme than that observed in the study under an assumption that the null hypothesis is true.
11. The correct answer is **C**. The null hypothesis is a statement of no effect or not association.
12. The correct answer is **F**. The probability of finding an effect or association when one exists is the power of the test. Power is defined as $1 - \beta$, where β is the probability of a type II error (failing to find an effect when one exists).
13. The correct answer is **D**. Declaring there is an effect or association when in fact there is not one is a type I error. (In this and the previous question, be sure you are distinguishing between the error and the *probability* of the error.)
14. The correct answer is **B**. We learned that the chi-square test is used to compare proportions. Although the data are presented in a 2 X 2 table, determining whether drug administration increases the incidence of ulcers is addressed by comparing the proportion of mice with ulcers in the drug and saline groups. <u>Another way to think about the chi-square test is that it is used to assess the association between two categorical variables</u>.
15. a. The correct answer is **B**. ANOVA is used to compare means across three or more groups.
    b. The correct answer is **D**. The chi-square test is used to compare proportions – here, the proportion of vitamin use between low and high fat consumption groups.
    c. The correct answer is **A**. A t-test is used to compare means between two groups.

16. The correct answer is **A**. Because the data are paired – baseline and one-year measures are obtained from the same subject – any test must account for this pairing. The only test we've learned that accounts for this is the paired t-test.
17. The correct answer is **A**. The null value for a ratio is 1. Since the 95% CI excludes the null value of 1, we can conclude that the finding is significant at $\alpha = 0.05$. This implies that the p-value is smaller than 0.05.
18. The correct answer is **D**. Recall that the CI width is inversely proportional to sample size – the larger the sample size, the narrower the interval. Since both studies report a similar value for the RR estimate, we can rule out choice A that study A overestimates the risk. Choice B is ruled out because one can never demonstrate causality in an observational study. We can rule out choice C – there is no reason to assume the results of study A are inaccurate (if they were, we'd see a disparate findings in the two studies). And choice E is incorrect – the 95% CI includes the null value so the p-value is greater than 0.05.
19. The correct answer is **B**. When confidence intervals do not overlap then there is a significant difference. However, when confidence intervals overlap, the difference may or may not be significant.
20. The correct answer is **B**. The 1.0 in the row corresponding to normal weight indicates that all ORs reported in the table are computed relative to the normal weight category. This rules out choices A and D. Choice C is incorrect because two of the three odds ratios reported in the obese class are not significant – the CIs contain 1. This leave choice B. To see that this is correct, note that the OR reported for the obese class 3 subjects is calculated relative to the normal weight. Since the OR is 4.98, and the 95% CI excludes 1, we can conclude that obese class 3 individuals have significantly higher odds of depression relative to normal weight individuals. Equivalently, normal weight individuals have significantly lower odds of depression relative to class 3 obese individuals.
21. The correct answer is **D**. Here we are comparing the two groups based on a difference, so the null value is 0. Since the 95% CI excludes 0, the difference is significant. Furthermore, from the table we can see that cholesterol levels were higher in the New Zealand group. Because we infer back to the target population, the correct answer is choice D.
22. The correct answer is **D**. The problem is asking for the Kaplan-Meier (or product-limit) estimate of survival. This is calculated by taking the cumulative product of the probability of being alive in each time interval.
23. The correct answer is **A**. The log-rank test compares two survival functions at all time points. It is not a test at a specific time point. Since the p-value for the log-rank test is 0.01, we conclude that survival is significantly increased in the treatment group relative to placebo. This is equivalent to concluding that the risk of death is significantly lower in the treatment group relative to placebo.
24. The correct answer is **B**. Since the log-rank test is significant, we know the HR must differ significantly from 1. Therefore, the corresponding 95% CI will exclude the null value of 1.0.