***STATISTICS***

Sample Test 4—ANSWERS

1. A relationship between two variables.
2. False
3. + (The longer you are in sun, the more severe the sunburn.)
4. – (If more people are unemployed, less money will be spent.)
5. + (The longer you exercise, the more calories you burn.)
6. 0 (probably not much relationship)
7. + (Bigger lakes can support more fish.)
8. – (Usually older people are less flexible than younger people.)
9. + (The longer since you ate, the hungrier you are.)
10. + (More wind turbines means more wind energy.)
11. 0 (There’s not much difference in gas purchases in different parts of a month.)
12. B (negative – D is impossible)
13. D (fairly strong positive)
14. C (positive – A is impossible
15. A (no real pattern)
16. D (strong negative)
17. A (perfect positive)
18. C (positive and negative cancel out)
19. C (weak positive)
20. B (nearly zero)
21. 15 (pairs of data)
22. r(15,.05) = .51
23.    , so r = .776
24. YES (because .776 > .51)
25. r2 = .602, so about 60%
26. 11
27. r(11,.01) = .73
28.   r = .65 (This was corrected from the original answer key.)
29. No (because (.65 < .73)
30. (-.6)2 = .36 or 36%
31. 100 – 36 = 64%
32. Here they give us that $r^{2}=.49$, so we need to take $\sqrt{.49}$, which is .7
33.  12.3 or about 12
34.  24
35. Here we know that $25=0.9\hat{x}+10.5$  $16.\overbar{1}$ or about 16
36. 4 x 2 (you don’t include the totals)
37.    P = .62085
38. No (because .62 > .05)
39.   P = .0106
40. Yes (because .0106 < .10)
41. Yes (because .0106 < .05)
42. No (because .0106 > .01)
43. C (.01 or 1%, because important medical research should have the lowest possible level of significance – that is, you don’t want to be wrong very often)

|  |  |
| --- | --- |
| L1 | L2 |
| 50 | .15\*200 = 30  |
| 75 | .35\*200 = 70 |
| 25 | .07\*200 = 14 |
| 20 | .13\*200 = 26 – This was wrong on the original answer sheet. |
| 30 | .30\*200 = 60 |

1. 4 (Since there are 5 categories, take 5 – 1 = 4)
2. P = 7.97 X 10-8 (The P-Value WAS correct as original posted.)
3. Yes (the p-value is much less than 10%.)
4. 9 (10 sports, so 10 – 1)
5. C (different to begin with)
6. E (highlighting one small thing that makes them look good)
7. B (refuse to answer—or lie—when asked personal questions)
8. F (the company has a vested interest in saying their product is beneficial)
9. D (college students at spring break aren’t necessarily like all Americans)
10. A (people act differently when they think they’re being watched or studied)
11. & 56.

**Standard deviation** $χ^{2}$ looks at whether data is too spread out (because the standard deviation is too large); **runs test** sees whether or not data is random; **Spearman’s r** looks at whether different rankings are different; **high power tests** are unlikely to overlook significant results (but may include false positives); **analysis of variance** compares more than 2 averages and standard deviations; **non-linear regression** looks at patterns that aren’t lines (like curved and cyclical data); **multiple regression analysis** uses many different factors to predict an outcome; **calculus based statistics** looks at constantly changing data and is used to find the areas under curves used for statistical probability

1. B (average and S.D. of 2 samples)
2. A (“The United Nations knows the S.D. for the world, and Denmark is likely to provide a large sample.)
3. B (as one thing increases, so does another)
4. B (As one thing goes up, the other goes down)
5. A (comparing percent from 2 different years)
6. C (His sample is more than 30 … and also the U.S. Bureau of Labor Statistics knows the S.D. for the whole country)
7. B (They organized the data in a table … this one would be a 2 x 4 matrix)
8. A (They’re comparing the average of 2 samples .. in this case it might actually be a 2-sample **z** test, since both samples are likely to be large, but that is not an option)
9. D (comparing % of men and % of women)
10. This will obviously vary depending on your project. You don’t need to know specific numbers, such as the p-value you got, but you should be able to answer parts a – d.