Principles and Applications of Biostatistics 1st edition Merrill Solutions Manual

SKU: 9781284225976-SOLUTIONS

Appendix A

Answers to Exercises

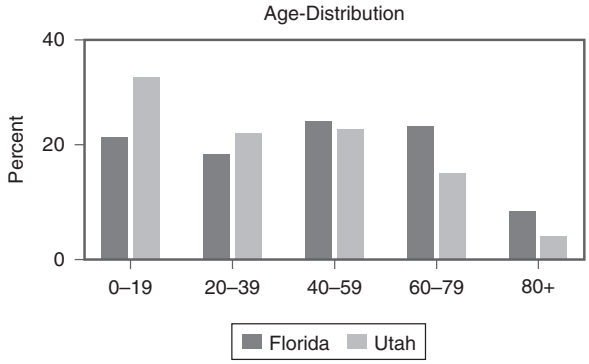
Chapter 1

1. Descriptive statistics, probability, inference, and statistical techniques. Descriptive statistics are measures of frequency, tendency, dispersion (variation), and position that represent either a sample or a population. Probability is used when discussing the chance or likelihood that a given event will occur. It is the foundation of statistical inference. Statistical inference is when conclusions are drawn about a population based on sample information taken from the population. Statistical techniques are analytic approaches that utilize various statistical methods to investigate research questions.
2. C, D, B, A
3. A, B, A, B, A
4. Discrete data represent things that are counted. Continuous data represent things that are measured. Continuous data are more precise and provide more information than discrete data.
5. Frequency (count, percent rate); central tendency (mean, median, mode); dispersion (standard deviation, range, interquartile range); position (percentile, rank)
6. Probability is a numerical description of how likely an event is to occur. Some applications of probability in biostatistics include risk assessment, sampling, evaluating the reliability of an estimator, validity of a diagnostic test, and the chance of survival over a certain duration.
7. Inferential statistics involves probability by drawing conclusions about the population based on sample data beyond what would be expected by random chance.
8. C, A, B, B, A
9. 8, 5.75, 5.34, 1, 16
10. The 95% confidence interval for the mean is 1.29, 10.21.

Chapter 2

1. A – Ratio
2. B – Proportion
3. C – Rate
4. Poorer recovery but better survival for Disease A than Disease B.
5. $34/63 \times 100 = 54.0$ per 100
6. D
7. A, B, A, B, C
8. The crude rate ratio is 1.63, meaning the death (mortality) rate in Florida is 1.63 times (63%) greater as the death rate in Utah.

9. The age distribution is younger in Utah; older in Florida.



10. Mortality rates per 100,000 are shown according to age group and area in the following table. Rates are higher in Florida in the age groups through 79 and higher in Utah in the age group 80 years and older.

	Florida	Utah
0–19	56.7	51.7
20–39	135.4	111.1
40–59	325.3	247.6
60–79	1224.5	1060.4
80+	6110.9	7463.7

11. The age-adjusted rate in Utah using the Florida population as the standard is 1011.1 per 100,000.
12. The crude rate in Florida is 957.2 per 100,000. Then, 957.2/1011.1 = 0.947, meaning after accounting for differences in the age-distribution, the mortality rate in Florida is 5.3% lower than in Utah.
13. Seven for ages 45–54, 12 for ages 55–64, 19 for ages 65–74, 31 for ages 75–84, and 52 for ages 85 and older.
14. SMR = 163/121 = 1.35; in Texas, 2014–2016, blacks experienced a 35% greater death rate from malnutrition.
15. Yes, if the age-distribution between groups are similar or if the rate is not dependent on age.
16. Age-adjustment of rates over time assumes that the age-specific rates have a similar trend over time; that the trend lines are approximately parallel.
17. The crude rate represents the actual risk of the event in the population. If the age-distribution differs between populations being compared or in a given population over time, then to make comparisons in rates that are not confounded by age, age-adjustment is necessary.
18. The standard population should be within a time frame in which rates are being compared. For example, if we are interested in evaluating changes in rates between 2000 and 2020, a standard population within this time frame should be used. If rates are being compared between groups, the standard should reflect the age distribution of at least one of these populations.
19. Arizona has a greater relative variability, as indicated by the coefficient of variation.

Analysis Variable : Utah						
Mean	Median	Variance	Std Dev	Range	Quartile Range	Coeff of Variation
424.7	429.5	3892.2	62.4	164.0	122.0	14.7

Analysis Variable : Arizona						
Mean	Median	Variance	Std Dev	Range	Quartile Range	Coeff of Variation
2319.6	2223.0	680442.7	824.9	2710.0	1179.0	35.6

20.

Analysis Variable : Utah						
Minimum	Lower Quartile	Median	Upper Quartile	Maximum	Range	Quartile Range
2231	2516	3003	3960	4035	1804	1444

21. No outliers were seen this week because no observations are < Q1 − (1.5 × IQR) = 350 or > Q3 + (1.5 × IQR) = 6126.
22. Line graph
23. A
24. The geometric mean (root of product) is appropriate for a set of positive numbers in which the values are meant to be multiplied or have an exponential distribution (e.g., a set of growth figures).
25. Yes, $\sqrt{.75 \times 8} = 2.4$ and $\sqrt{.8 \times 5} = 2.0$.
26. The variance is a measure of scatter (dispersion) of observations around the mean. A sample is a subset of the population. Hence, the range of sample data will be smaller than the range of population data; the mean of the deviations about the mean will be smaller for sample data than it will be for population data. To adjust the variance up to where it better estimates the population variance, we divide by $n-1$ instead of N.
27. Degrees of freedom of an estimate is the number of independent pieces of information used to obtain the estimate.
28. When the frequency distribution of the populaton is symmetrical and bell-shaped (normal).
29. 68%, 95%, 99.7%
30. 0.3%

Chapter 3

1. If an experiment is repeated n times under the same conditions, and if the event A occurs m times, then as n grows larger, the ratio m/n approaches a fixed limit that is the probability of A. $P(A) = m/n$.
2. Classical probability measures the likelihood of something happening, with each outcome being equally likely to occur and the outcomes are mutually exclusive.
3. Subjective probability involves using one's own experience, judgment, or opinion to find probabilities.
4. B, B, A, C
5. Mutually exclusive events involve two or more events for which the occurrence of one event precludes the occurrence of the other(s). Independent events are events whose occurrence or outcome has no effect on the probability of the other. There is a link between mutually exclusive events in that they cannot both happen at once. However, there is no link between independent events.
6. When the events being considered are mutually exclusive.
7. $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$

8. $P(A \text{ or } B) = P(A) + P(B) = 0.15 + 0.10 = 0.25$
9. When the events are independent: $P(A \text{ and } B) = P(A)P(B)$
10. These two events are independent if the joint probability equals the product of the marginal probabilities.

    $$P(\text{Man and No}) = 0.410 \neq P(\text{Man}) \times P(\text{No}) = 0.475 \times 0.833 = 0.396,$$

    which indicates that the events are not independent.

    Another way to look at this is if $P(\text{Man} \mid \text{No}) = \dfrac{P(\text{Man and No})}{P(\text{No})} = P(\text{Man})$,

    then the two events are independent. This approach also shows the events are not independent: $P(\text{Man} \mid \text{No}) = 0.492 \neq P(\text{Man}) = 0.475$.
11. 0.966, 0.033, 0.001
12. Increases exponentially with age, from 0.012 (1.2%) for ages < 15 to 0.045 (4.5%) in ages 40–44, and then 0.128 (12.8%) in ages 45–54.
13. No, like the situation with twins, the proportion of women having triplets or higher increases exponentially with age. Beware if you get pregnant at age 45 or older!
14. 95.93%
15. 3.93%
16. 0.13%
17. 0.009
18. 0.033
19. Bayes' Theorem relates the probability of the occurrence of an event to the presence or nonpresence of an associated event; that is, it shows that if we know the conditional probability of B given A and $P(B)$ and $P(A)$, we can identify the conditional probability of A given B.
20. Sensitivity = 285/300 = 0.95; of those abused, 95% are identified as such by a physical exam. Sensitivity = 8730/9700 = 0.90; of those who are not abused, 90% are identified as such by a physical exam.
21. $PV^+ = 0.227$; of those with a positive physical exam for abuse, 22.7% have been abused. $PV^- = 0.998$; of those with a negative physical exam for abuse, 99.8% have not been abused.
22. Refer to Table 3.4.
23. Cluster sampling.
24. Stratified sampling.
25. Do not use systematic sampling if the data exhibits patterns. On the other hand, if it does not exhibit patterns and the risk of data manipulation by a researcher is low, it may be a less expensive, straightforward approach.
26. Selection bias.
27. Use "=RANDBETWEEN(1,500)" in a given cell and copy and paste it into the number of cells of your desired sample size.

Chapter 4

1. A random variable is usually written as X, with possible values that are numerical outcomes of a random phenomenon of interest.
2. Discrete and continuous random variables are similar in that they may produce interval or ratio data. They are different in that discrete data have a finite number of possible values, but continuous data have a theoretically infinite number of possible values.