ESSEXCOUNTY MATH LEAGUE
STATISTICS SOLUTIONS
MAY 22, 2013

1. An eight-sided die is numbered 1-8, a twelve sided die is numbered 1-12, and a twenty-sided die is numbered 1-20. All three dice are rolled. Assuming all dice are fair, what is the probability of rolling a 5 on at least two of the dice?

   A) 0.0115 B) 0.0193 C) 0.0198 D) 0.0208

   Solution: (C) Let F_1, F_2, F_3, be the event of rolling a five on the eight-sided die, twelve-sided die, twenty-sided die, respectively. Then $P(\text{at least two 5s}) = P(F_1 \cap F_2 \cap \overline{F}_3) + P(F_1 \cap \overline{F}_2 \cap F_3) + P(\overline{F}_1 \cap F_2 \cap F_3) + P(F_1 \cap F_2 \cap F_3) = P(F_1)P(F_2)P(\overline{F}_3) + P(F_1)P(\overline{F}_2)P(F_3) + P(\overline{F}_1)P(F_2)P(F_3) + P(F_1)P(F_2)P(F_3) = \frac{1}{8}\cdot\frac{1}{12}\cdot\frac{19}{20} + \frac{1}{8}\cdot\frac{11}{12}\cdot\frac{1}{20} + \frac{7}{8}\cdot\frac{1}{12}\cdot\frac{1}{20} + \frac{1}{8}\cdot\frac{1}{12}\cdot\frac{1}{20} = \frac{19+11+7+1}{1920} = \frac{38}{1920} = 0.0198$.

2. Which one of the following distributions is <u>not</u> symmetrical with respect to the mean.

   A) Normal B) t C) χ^2 D) Uniform
   E) All are symmetrical with respect to the mean

   Solution: (C) The Normal, t, and Uniform distributions are all symmetrical with respect to the mean. The χ^2 distribution is not symmetrical with respect to the mean.

3. A data set consists of scores with one outlier. Which of the following will change the most if the outlier is removed from the data.

   A) Mean B) Median C) Mode D) Midrange

   Solution: (D) The mean will be affected slightly since all scores are involved in the calculation. The median may be affected slightly, since removing the outlier will shift the middle score(s) slightly. The mode will not be affected at all. Since the midrange involves only the two extreme scores, it will be affected the most.

4. Which of the following are examples of a discrete random variable.

   I) Shoe sizes as labeled on shoe boxes
   II) The length of yarn scraps
   III) The number of apples picked in an orchard
   IV) The time it takes commuters to get to work

   A) I,III only B) II,III only C) III only D) III,IV only E) IV only

   Solution: (A) Both I and III consist of values that can be listed, including shoe sizes, which include fractions. So these are discrete random variables. Both II and IV consist of values that include a range of real numbers. So these are continuous random variables.

5. A coin is tossed five times. Let H_2 be the event that heads are tossed on the second toss, T_5 be the event that tails are tossed on the fifth toss, E_1 be the event that heads are tossed on the first four tosses and tails are tossed on the fifth toss, and E_2 be the event that heads are tossed on the first toss and tails are tossed on the last four tosses. Which of the following pairs of events are mutually exclusive?

I) H_2 and T_5 II) H_2 and E_1 III) H_2 and E_2 IV) E_1 and E_2

A) I only B) II,III only C) I,III,IV only D) III only E) III,IV only

Solution: (E) It is possible to toss heads on the second toss and toss tails on the fifth toss, so H_2 and T_5 are not mutually exclusive. Since H_2 includes the possibility of E_1 occurring, H_2 and E_1 are not mutually exclusive. Since it is impossible for heads to be tossed on the second toss when E_2 occurs, H_2 and E_2 are mutually exclusive. Since, for example, heads are tossed on the second toss for E_1 to occur and tails are tossed for E_2 to occur, E_1 and E_2 are mutually exclusive.

6. A recently discovered species of bats is claimed to have a mean of 510.3 grams. A random sample of 25 bats were weighed and found to have a mean of 524.7 grams and a standard deviation of 11.7 grams. A hypothesis test is conducted on the claim at the 0.05 significance level. Find the appropriate test statistic.

A) –6.15 B) –1.23 C) 1.23 D) 6.15

Solution: (D) This is a test of means where the population standard deviation is not known, so this involves the t distribution. The population mean is claimed to be 510.3, the sample mean is 524.7, the sample standard deviation is 11.7, and the sample size is 25. The appropriate test statistic is

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{524.7 - 510.3}{11.7/\sqrt{25}} = 6.15.$$

7. For the problem above, find the appropriate critical value(s). Note: indicate only the positive critical value if there are two.

A) 1.645 B) 1.708 C) 1.711 D) 1.960 E) 2.064

Solution: (E) This is a two-tail test, so there are two critical values, and the area of the critical region to the right of the positive critical value is 0.025. There are 24 degrees of freedom. So from Table B, the positive critical value is 2.064.

8. A given set of data has a standard deviation of 3. All the scores are multiplied by –2 and then added by 4. What is the new standard deviation?

A) –6 B) –2 C) 6 D) 10
E) Not enough information to determine

Solution: (C) When multiplying the scores of the data by –2, the standard deviation changes by a factor of $|-2| = 2$. Added the same number to each score does not change the standard deviation further. So the standard deviation of the new data is 6.

9. A fair coin and a fair die are tossed. If heads are tossed or a 3 is rolled, Tom wins $10 from Bill. If tails are tossed and an even number is rolled, Bill wins $20 from Tom. Otherwise, no money is exchanged. What is Tom's expectation for this game?

A) −$2.50 B) $0 C) $0.83 D) $1.67

Solution: (C) There are 12 equally likely possibilities when tossing a fair coin and a fair die. Tom wins $10 when the event {H1,H2,H3,T3,H4,H5,H6} occurs, which has probability 7/12. Tom loses $20 when the event {T2,T4,T6} occurs, which has probability of 1/4. Tom's expectation is $10 \cdot \frac{7}{12} + (-20) \cdot \frac{1}{4} = \$0.83$.

10. The weights of widgets are normally distributed with a mean of 23.4 ounces and a standard deviation of 4.7 ounces. A random sample of widgets is chosen. What is the probability that the mean weight is less than 25.1 ounces?

A) 0.6406 B) 0.9554 C) 0.9999 D) Not enough information

Solution: (D) Since the probability to be determined involves the mean weight of the sample, in order to find the probability, the sample size would be needed. There is not enough information in this problem to determine the sample size.

11. A cola bottling plant claims that the mean amount of cola in their twelve ounce labeled bottles is more than 11.8 ounces. A hypothesis test is conducted on the plant's claim. The company will retain its current bottle filling procedure if the claim is supported. Which of the following represents a Type I error?

A) Retaining the bottle filling procedure when the actual mean is at most 11.8 ounces.
B) Changing the bottle filling procedure when the actual mean is not 12.0 ounces.
C) Changing the bottle filling procedure when the actual mean is more than 11.8 ounces.
D) Changing the bottle filling procedure when the actual mean is at most 11.8 ounces.
E) Retaining the bottle filling procedure when the actual mean is more than 11.8 ounces.

Solution: (A) A Type I error occurs if the null hypothesis is rejected when it is true, or equivalently, if the alternate hypothesis is supported when it is false. The alternate hypothesis for this situation is the mean amount of cola is more than 11.8 ounces. If true, then the plant would retain the bottle filling procedure. However, with a Type I error, the procedure would be retained even though the statement "more than 11.8 ounces" is false., i.e., the actual mean is at most 11.8 ounces.

12. One wheel is evenly divided into three sections numbered 0,1,2. A second wheel is evenly divided into three sections numbered 1,2,3. Both wheels are spun. Let X be the product of the numbers on the wheels. Find the standard deviation of X.

A) 0.7 B) 1.9 C) 2.8 D) 3.3

Solution: (B) The elements of the sample space are (0,1), (0,2), (0,3), (1,1), (1,2), (1,3), (2,1), (2,2), (2,3). The respective values of X are 0, 0, 0, 1, 2, 3, 2, 4, 6. So the probability distribution is

x	0	1	2	3	4	6
$p(x)$	1/3	1/9	2/9	1/9	1/9	1/9

$\mu = E(X) = \sum x \cdot p(x) = 0 \cdot \frac{1}{3} + 1 \cdot \frac{1}{9} + 2 \cdot \frac{2}{9} + 3 \cdot \frac{1}{9} + 4 \cdot \frac{1}{9} + 6 \cdot \frac{1}{9} = 2.$

$Var(X) = \sum (x-\mu)^2 p(x) = (0-2)^2 \cdot \frac{1}{3} + (1-2)^2 \cdot \frac{1}{9} + (2-2)^2 \cdot \frac{2}{9} + (3-2)^2 \cdot \frac{1}{9} + (4-2)^2 \cdot \frac{1}{9} + (6-2)^2 \cdot \frac{1}{9} = \frac{34}{9}.$

$\sigma(X) = \sqrt{\frac{34}{9}} = 1.9.$

13. A fair six-sided die is rolled. If a 1 or 2 is rolled, a fair coin is flipped twice. Otherwise, a fair coin is flipped three times. Find the probability of rolling a 1 or 2 given that exactly one tail is tossed.

A) 1/4 B) 2/5 C) 1/2 D) 4/7

Solution: (B) Let A be the event a 1 or 2 is rolled on the die, and let B be the event of tossing one tail. When two coins are tossed, the possibilities are HH, HT, TH, TT, so $P(B|A)=\frac{1}{2}$. When three coins are tossed, the possibilities are HHH, HHT, HTH, HTT, THH, THT, TTH, TTT, so $P(B|\bar{A})=\frac{3}{8}$.

$$P(A|B)=\frac{P(A\cap B)}{P(B)}=\frac{P(A\cap B)}{P(A\cap B)+P(\bar{A}\cap B)}=\frac{P(A)P(B|A)}{P(A)P(B|A)+P(\bar{A})P(B|\bar{A})}=\frac{(1/3)(1/2)}{(1/3)(1/2)+(2/3)(3/8)}=$$

$$\frac{1/6}{1/6+1/4}=\frac{2/12}{5/12}=\frac{2}{5}.$$

14. All paired data points (x,y) satisfy the equation $2x + 3y = 36$. Which of the following is the correlation coefficient?

A) –1 B) –2/3 C) 0 D) 2/3 E) 1

Solution: (A) Because the data points all fall on the line, there is perfect linear correlation. The equation of the regression line is $\hat{y}=-\frac{2}{3}x+12$, indicating negative correlation. Thus, the correlation coefficient is –1.

15. The weight of turkeys on a poultry farm is normally distributed with a mean of 21.4 pounds. It was found that 9% of the turkeys weight more than 24.1 pounds. Find the percentage of turkeys that weigh less than 22.3 pounds.

A) 0.6368 B) 0.6736 C) 0.8159 D) 0.9641
F) Not enough information

Solution: (B) Let X be the normal random variable representing the weight of turkeys. The mean is 21.4. Let σ be the standard deviation. From the information given, $P(X>24.1)=P\left(Z>\frac{24.1-21.4}{\sigma}\right)=$

$P\left(Z>\frac{2.7}{\sigma}\right)=0.09\rightarrow P\left(Z<\frac{2.7}{\sigma}\right)=1-0.09=0.91$. From Table A, $P(Z<1.34)=0.91\rightarrow\frac{2.7}{\sigma}=1.34\rightarrow$

$\sigma=\frac{2.7}{1.34}=2.0$. $P(X<22.3)=P\left(Z<\frac{22.3-21.4}{2.0}\right)=P(Z<0.45)=0.6736$.

16. A company sells boxes of candy consisting of five flavors. A consumer group tests the claim that the proportion of flavors are equal at the 0.05 significance level by testing a random sample of 30 candies. Find the appropriate critical value for this test.

A) 9.49 B) 11.07 C) 42.56 D) 43.77

Solution: (A) This is a goodness of fit (multinomial experiment) test, is a right tail test that uses the χ^2 distribution. The degrees of freedom is one fewer than the number of categories, which is 4. From Table C, the critical value is 9.49

17. A car rental agency has 225 reservations for Friday, and 190 cars available. Current studies have shown that 80% of customers keep the reservations. Use normal approximation to approximate the probability that all the customers who keep the reservations on Friday will get a car.

A) 0.9332 B) 0.9429 C) 0.9525 D) 0.9599

Solution: (D) Let X be the binomial random variable representing the number of customers that keep their reservations. Let $\tilde{X}$ be the normal approximation of X. The mean of X is $\mu = np = 225(0.8) = 180$, and the standard deviation of X is $\sigma = \sqrt{np(1-p)} = \sqrt{225(0.8)(0.2)} = 6$. In order for there to be enough cars on Friday, at most 190 of the customers must keep the reservations. $P(X \le 190) \approx P(\tilde{X} < 190.5) =$

$$P\left(Z < \frac{190.5-180}{6}\right) = P(Z < 1.75) = 0.9599.$$

18. While doing research on data, Sue calculated the mean and standard deviation of a sample to be 17.1 and 4.4, respectively. When reviewing her calculation, Sue discovered that one of the scores was erroneously recorded as 12 was really 21. Which of the following statements is true?

A) The actual sample standard deviation is less than 4.4
B) The actual sample standard deviation is 4.4
C) The actual sample standard deviation is greater than 4.4
D) Not enough information to make one of the above conclusions.

Solution: (A) The score of 12 is 5.1 units away from the mean. The score of 21 is closer to the mean of 17.1. In fact, the actual mean is slightly higher. Thus, the standard deviation will decrease.

19. Because the probability of having a positive blood test for disease H from one person is 0.001, a lab will combine 40 samples to test. If the test is negative, that means that all 40 samples are negative for disease H. Otherwise, the tests will have to be done on the individual samples. If a lab conducts such a test, what is the probability that the lab will have to test all 40 samples?

A) 0.000761 B) 0.0250 C) 0.0384 D) 0.0392

Solution: (D) Let X be the binomial random variable that represents the number of positive tests. We have $n = 40$ and $p = 0.001$. The lab will have to test all 40 samples if X is greater than 0. $P(X > 0) =$

$$1 - P(X = 0) = 1 - \left[\binom{40}{0}(0.001)^0(0.999)^{40}\right] = 1 - 0.9608 = 0.0392.$$

20. A company is developing a vending machine that accepts quarters. The machine will accept a quarter provided the quarter weighs less than 5.73 grams. A study was conducted in which random samples of 50 quarters were weighed, and it was found that the probability that the mean weight of a sample is less than 5.73 grams is 0.998. Which of the following statements is not true based on the above information?

A) The mean weight of a quarter is less than 5.73 grams.
B) About 99.8% of quarters would be accepted by the machine.
C) More than half of the quarters would be accepted by the machine.
D) It is not known what percentage of coins would be accepted by the machine.
E) All of the statements are true.

Solution: (B) The sample mean of 50 quarters is less than 5.73. Since the sample mean is equal to the mean of the individual quarters, statement A is true. As such, more than half of the quarters would be accepted, so statement C is true. Since the standard deviation cannot be determined by the above information, statement D is true. While there is a 99.8% chance that the sample mean is less than 5.73 grams, a smaller percentage (but still more than half) of individual quarters would be accepted by the machine (as per Central Limit Theorem). Thus, statement B is not true.

21. Consider the incomplete contingency table below. The numbers in the total rows or columns represent the actual total for the given row or column. The numbers in the Yes-A and No-C cells represent the expected numbers (rounded off) if the row and column variables are independent. Find the grand total.

	Yes	No	Not Sure	Total
A	103.59			x
B				263
C		164.21		315
Total	350	y	43	t

Solution: (821) We have the following equations relating x and t. $x+263+315=t \rightarrow x=t-578$ and $\frac{350x}{t}=103.59 \rightarrow x=\frac{103.59t}{350}$. Setting the equations equal to each other, we get $t-578=\frac{103.59t}{350} \rightarrow t-\frac{103.59t}{350}=578 \rightarrow \left(1-\frac{103.59}{350}\right)t=578 \rightarrow \frac{246.41}{350}t=578 \rightarrow t=821$. Equations relating y and t can also be used.