MSCA 31000 - Introduction to Statistical Concepts

Chapter 5

Q1: (a) What is the probability of rolling a pair of dice and obtaining a total score of 9 or more?
(b) What is the probability of rolling a pair of dice and obtaining a total score of 7?

Answer:
(a) P(9 or 10 or 11 or 12) = 10/36 = 0.278
(b) P(7) = 6/36 = 0.167

Q3: A card is drawn at random from a deck. (a) What is the probability that it is an ace or a king?
(b) What is the probability that it is either a red card or a black card?

Answer:
(a) P(Ace or King) = 0.1538462
(b) P(Red or Black card) = 1

Q5: A fair coin is flipped 9 times. What is the probability of getting exactly 6 heads?

Answer: 0.1640625

Q7: You flip a coin three times. (a) What is the probability of getting heads on only one of your flips? (b) What is the probability of getting heads on at least one flip?

Answer:
(a) 0.375
(b) 0.875

Q9: A jar contains 10 blue marbles, 5 red marbles, 4 green marbles, and 1 yellow marble. Two marbles are chosen (without replacement).

(a) What is the probability that one will be green and the other red? (b) What is the probability that one will be blue and the other yellow?

Answer:
(a) P(1 Green and 1 Red) = P(Green first and then Red without replacement) + P(Red first and then Green without replacement) = **0.105**
(b) P(1 Blue and 1 Yellow) = P(Blue first and then Yellow without replacement) + P(Yellow first and then Blue without replacement) = **0.0526**

Chapter 7

Q2: (a) What are the mean and standard deviation of the standard normal distribution? (b) What would be the mean and standard deviation of a distribution created by multiplying the standard normal distribution by 8 and then adding 75?

Answer:
- (a) For a normal distribution, mean = 0, std dev = 1.
- (b) For Linear Transformation (Y = bX + A), where b = 8 and A = 75
    - mean = bμ+A = 8*0 + 75 = **75**
    - std dev = bσ = 8 * 1 = **8**

Q4: (a) What proportion of a normal distribution is within one standard deviation of the mean? (b) What proportion is more than 2.0 standard deviations from the mean? (c) What proportion is between 1.25 and 2.1 standard deviations above the mean?

Answer:
- (a) 0.68
- (b) 0.0455
- (c) 0.088

Q6: Assume a normal distribution with a mean of 70 and a standard deviation of 12. What limits would include the middle 65% of the cases?

Answer: Plugging relevant numbers into http://onlinestatbook.com/2/calculators/normal.html, we get the range as **58.786 to 81.214**

Q8: Assume the speed of vehicles along a stretch of I-10 has an approximately normal distribution with a mean of 71 mph and a standard deviation of 8 mph.
- a. The current speed limit is 65 mph. What is the proportion of vehicles less than or equal to the speed limit?
- b. What proportion of the vehicles would be going less than 50 mph?
- c. A new speed limit will be initiated such that approximately 10% of vehicles will be over the speed limit. What is the new speed limit based on this criterion?
- d. In what way do you think the actual distribution of speeds differs from a normal distribution?

Answer: Plugging relevant numbers into http://onlinestatbook.com/2/calculators/normal.html
- a. 0.2266
- b. 0.0043
- c. 81.254 mph
- d. I think that the distribution will not be a standard normal distribution and will be negatively skewed. There are usually minimum and maximum speed levels, and most people drive close to the upper limit which will result in the skew.

Q10: You want to use the normal distribution to approximate the binomial distribution. Explain what you need to do to find the probability of obtaining exactly 7 heads out of 12 flips.

Answer: For the binomial distribution with 12 flips,
- mean $\mu = N\pi = 12 * 0.5 = 6$
- variance $\sigma^2 = N\pi(1-\pi) = 12 * .5 * .5 = 3$

A total of 7 heads is $(7 - 6)/3 = 0.67$ standard deviations above the mean of the distribution. The question then is, "What is the probability of getting a value exactly 0.67 standard deviations above the mean?" But this will give us 0 since the probability of a discrete value on a normal distribution is 0.

As a result, we will expand the range from 6.5 to 7.5 and consider that the area under the curve will give us the probability of obtaining 7 heads.

Plugging this value into the calculator results in the probability value: **0.1932**.

Chapter 9

Q1: A population has a mean of 50 and a standard deviation of 6. (a) What are the mean and standard deviation of the sampling distribution of the mean for N = 16? (b) What are the mean and standard deviation of the sampling distribution of the mean for N = 20?

Answer:
- (a) For N = 16,
  - Mean of sampling distribution = mean of population = **50**
  - Std dev of sampling distribution = Std dev of population / sqrt(N) = 6/sqrt(16) = **1.5**
- (a) For N = 20,
  - Mean of sampling distribution = mean of population = **50**
  - Std dev of sampling distribution = Std dev of population / sqrt(N) = 6/sqrt(20) = **1.34**

Q3: What term refers to the standard deviation of the sampling distribution?

Answer: Standard error of the mean.

Q5: A questionnaire is developed to assess women's and men's attitudes toward using animals in research. One question asks whether animal research is wrong and is answered on a 7-point scale. Assume that in the population, the mean for women is 5, the mean for men is 4, and the standard deviation for both groups is 1.5. Assume the scores are normally distributed. If 12 women and 12 men are selected randomly, what is the probability that the mean of the women will be more than 1.5 points higher than the mean of the men?

Answer: 0.21

Q7: If numerous samples of N = 15 are taken from a uniform distribution and a relative frequency distribution of the means is drawn, what would be the shape of the frequency distribution?

Answer: N = 15 is a relatively large sample size. With fewer samples (say 1,000), the frequency distribution will approach a normal distribution with a minor positive skew. With several samples (say >50,000), the frequency distribution will be a normal distribution.

Q9: What is the shape of the sampling distribution of r? In what way does the shape depend on the size of the population correlation?

Answer: Sampling distribution of r is negatively skewed (unless r = 0). The greater the value of ρ (population correlation), the more pronounced the skew.

Q11: A variable is normally distributed with a mean of 120 and a standard deviation of 5. Four scores are randomly sampled. What is the probability that the mean of the four scores is above 127?

Answer:

Mean of sample = mean of population = 120.
Std dev of sample = Std dev of population / sqrt(N) = 5 / sqrt(4) = 2.5

Feeding this data into the calculator and looking for probability of area above 127.

Results in probability value = **0.0026**

Q13: The mean GPA for students in School A is 3.0; the mean GPA for students in School B is 2.8. The standard deviation in both schools is 0.25. The GPAs of both schools are normally distributed. If 9 students are randomly sampled from each school, what is the probability that:

(a) the sample mean for School A will exceed that of School B by 0.5 or more?
(b) the sample mean for School B will be greater than the sample mean for School A?

Answer: Using formulae for Sampling Distribution of Difference Between Means

     a)   mean difference of sample means = mean$_{School A}$ - mean$_{School B}$ = 3 - 2.8 = 0.2
          Std Dev of samples = sqrt(2 * (std dev)^2 / N) = 0.118

          Feeding into calculator, probability of Sample Mean for School A exceeding that of School B = **0.0055**

     b)   mean of sample = mean$_{School B}$ - mean$_{School A}$= 2.8 - 3 = -0.2
          Std Dev difference of sample = sqrt(2 * (std dev)^2 / N) = 0.118

          The mean + std dev for B only exceeds the sample mean +/ std dev after **4 standard deviations**. Plugging this into the calculator, the probability that sample mean for B will exceed sample mean for A is **0.000032**

Q15: When solving problems where you need the sampling distribution of r, what is the reason for converting from r to z'?

Answer: Since sampling distribution of r is not a normal distribution, we need to transform it into a normal distribution for probability calculations.

Q17: True/false: The standard error of the mean is smaller when N = 20 than when N = 10.

Answer: True.

Q19: True/false: You choose 20 students from the population and calculate the mean of their test scores. You repeat this process 100 times and plot the distribution of the means. In this case, the sample size is 100.

Answer: False. The sample size is 20 and number of samples is 100.

Q21: True/false: The median has a sampling distribution.

Answer: True. All statistics have a sampling distribution.