SSIE660: Stochastic Systems

Quiz 2

Thursday, Oct 17, 2016

Name: Key

1. Closed book
2. Calculator allowed.

1. Circle T or F. Briefly explain if you choose F. (2 points each, total 10 points)

   (a) (T F) If the right whisker is longer than the left whisker in a box plot, then the data is negatively skewed.

   F: Positively skewed.

   (b) (T F) A child's grade (e.g., 1st grade in an elementary school) can be classified as a discrete variable.

   F: Categorical variable.

   (c) (T F) If you double each entry on a list, that doubles the standard deviation.

   T

   (d) (T F) $P(Z < 1.23) = P(Z < -1.23)$

   F: $1 - P(Z < 1.23) = P(Z < -1.23)$

   (e) (T F) The following terms are equivalent to each other: 1) Random Variable and 2) Random Variate.

   F: Random variable is a function that associates an entry in the sample space with a numerical value.

2. Consider the following list.

$$2, 40, 52, 55, 60, 70, 75, 85, 85, 90, 90, 92, 93, 94, 95, 98, 100, 115, 123, 215$$

   (a) Find the five-number summary. (3 points)

   Min = 2, Q1 =65 , Median = 90, Q3=(95+98)/2=96.5, Max = 215

(b) Identify any mild and/or extreme outlier. (3 points)

f_s = Q3-Q1 = 96.5 - 65 = 31.5, 1.5 f_s= 47.25, 3f_s= 94.5

Q1 - 1.5 f_s= 65- 47.25 =17.75

Q1 - 3f_s= -29.5

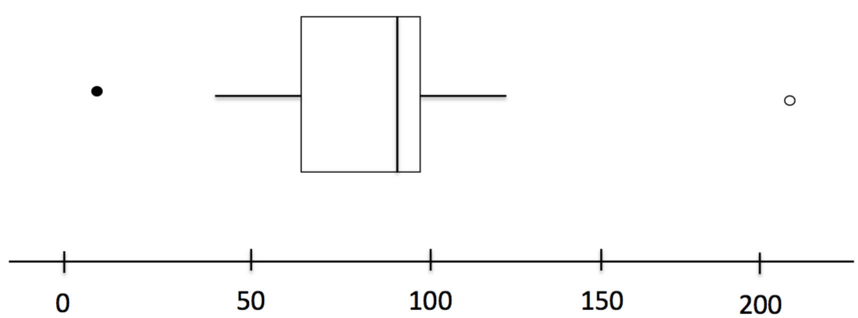
Q3 + 1.5 f_s= 96.5+ 47.25 =143.75

Q3 + 3f_s= 191

Mild outlier: 2

Extreme outlier: 215

(c) Draw a box plot showing mild and extreme outliers. (3 points)



3. Scores on the SAT Verbal test in recent years follow approximately the $N(500, 100)$ distribution. (4 points each, total 8 points)

   (a) How high must a student score in order to place in the top 10% of all students taking the SAT?

   Looking at the Z-score table, the entry closest to 0.9 is 0.8997. The corresponding z score is 1.28. Let x denote the student's score, then $z = \frac{x-500}{100}$ such that $x = 500 + 100 * 1.28 = 628$.

   (b) Calculate the percentile of a student whose score is 635.

   The z score is $z = \frac{635-500}{100} = 1.35$. The percentile is the area to the left of the z-score in the standard normal distribution. Looking at the Z score table, we find the area is 0.9115, such that the percentile is 91.

4. Find each of the following probabilities. (2 points each)

   (a) $P(Z < -0.23)$

   0.4090

   (b) $P(Z > 1.93)$

   $P(Z > 1.93) = 1 - P(Z < 1.93) = 1 - 0.9732 = 0.0268$

   (c) $P(0.65 < Z < 2.10)$

   $P(0.65 < Z < 2.10) = P(Z < 2.10) - P(Z < 0.65) = 0.9821 - 0.7422 = 0.2399$

5. The random variable X has probability density function

$$f(x) = \begin{cases} k(x-2)^2 & 1 < x < 2 \\ 0 & \text{elsewhere} \end{cases}$$

   (a) Find k. (4 points)

$$\int_1^2 k(x-2)^2 dx = k\int_1^2 (x^2 - 4x + 4)dx = k\left[\frac{1}{3}x^3 - 2x^2 + 4x\right]_1^2 = 1$$

$$= k\left[\frac{1}{3}(8-1) - 2(4-1) + 4(1)\right]_1^2 = 1$$

   By solving the above equation, we get $k = 3$

(b) Calculate $P\left[X > \frac{3}{2}\right]$. (4 points)

$$\int_{3/2}^{2} 3(x-2)^2 dx = \frac{1}{8}$$

(c) Find $E[X]$. (4 points)

$$\int_{1}^{2} 3x(x-2)^2 dx = \frac{5}{4}$$

6. An urn contains five red, three orange, and two blue balls. Three balls are randomly selected.

   (a) Let X represent the number of orange balls selected. What are the possible values of X? (3 points)

   $X = 0, 1, 2, 3$

   (b) Calculate $P[X = 1]$. (3 points)

   $$\frac{({}_3C_1)({}_7C_2)}{{}_{10}C_3}$$

7. Briefly explain the reason why we need to use $n-1$ as a denominator, instead of n, when we calculate a sample variance. (3 points)

   Sample variance is used as an estimate of the population variance. We could define the sample variance as the average squared deviation of the sample x_is about the population mean μ. However, the value of μ is almost never known, so the sum of squared deviation about $\bar{x}$ must be used. But the x_is tend to be closer to their average $\bar{x}$ than to the population average μ, so as to compensate for this the divisor $n-1$ is used.

8. Suppose that we toss 2 fair dice. Let E denote the event that the sum of the dice is 6 and F denote the event that the first die equals 4. Are E and F independent? or dependent? Show it. (5 points)

   Dependent. $P(E) = \frac{5}{36}$, $P(F) = \frac{6}{36}$ Therefore $P(E) \cdot P(F) \neq P(EF) = \frac{1}{36}$

9. Suppose the probability mass function of X is given by

$$p(x) = \begin{cases} .1, & x = 1 \\ .3, & x = 2 \\ .5, & x = 7 \\ .1, & x = 9 \end{cases}$$

   Find the cumulative distribution function of X and draw the graph of cdf of X. (5 points)

$$F(x) = \begin{cases} 0 & x < 1 \\ .1, & 1 \le x < 2 \\ .4, & 2 \le x < 7 \\ .9, & 7 \le x < 9 \\ 1, & 9 \le x \end{cases}$$

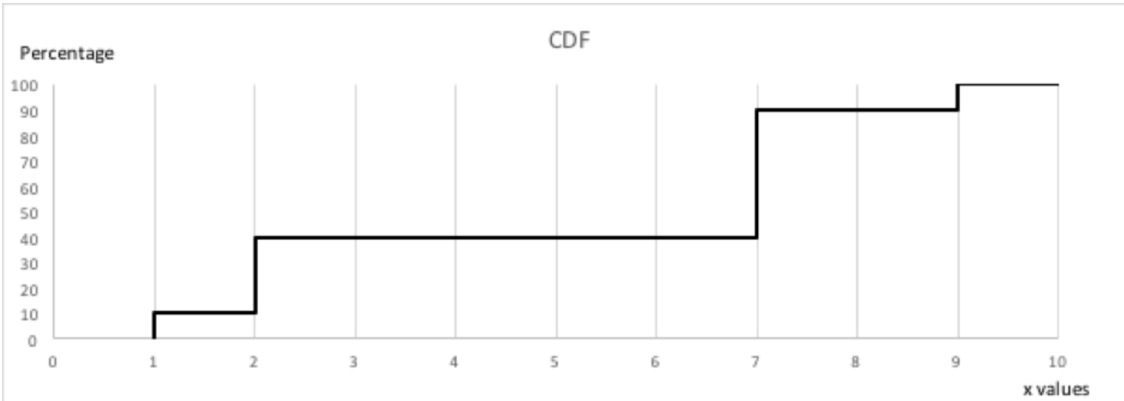


Figure1: CDF

10. A system composed of n separate components is said to be a parallel system if it functions when at least one of the components functions. For such a system, if component i, independent of other components, functions with probability $p_i, i = 1, 2, ..., n$, what is the probability that system functions? (5 points)

    $1 - (1 - p_1)(1 - p_2) \cdots (1 - p_n) = 1 - \prod_{i=1}^{n}(1 - p_i)$

11. A laboratory blood test is 95 percent effective in detecting a certain disease when it is, in fact, present. However, the test also yields a "false positive" result for 1 percent of the healthy persons tested. (That is, if a healthy person is tested, then, with probability 0.01, the test result will imply he or she has the disease.) If 0.5 percent of the population actually has the disease, what is the probability a person has the disease given the the test result is positive? (10 points)

    Let D be the event that the tested person has the disease and E the event that the test result is positive. The desired probability $P(D|E)$ is obtained by

$$\begin{aligned} P(D|E) &= \frac{P(DE)}{P(E)} \\ &= \frac{P(E|D)P(D)}{P(E|D)P(D) + P(E|D^c)P(D^c)} \\ &= \frac{(.95)(.005)}{(.95)(.005) + (.01)(.995)} \\ &= \frac{95}{294} \approx .323 \end{aligned}$$

12. Buses arrive at a specified stop at 15-minute intervals starting at 7 A.M. That is, they arrive at 7, 7:15, 7:30, 7:45 and so on. If a passenger arrives at the stop at a time that is uniformly distributed between 7 and 7:30, find the probability that he waits less than 5 minutes for a bus. (8 points)

$$P[10 < X < 15] + P[25 < X < 30] = \int_{10}^{15} \frac{1}{30} dx + \int_{25}^{30} \frac{1}{30} dx = \frac{1}{3}$$

13. The ideal size of a first-year class at a particular college is 500 students. The college, knowing from past experience that on the average only 30 percent of those accepted for admission will actually attend, uses a policy of approving the applications of 1500 students. Compute the probability that more than 500 first-year students attend this college. (13 points)

$$P\{X \geq 501\} \approx P\{X \geq 500.5\} = P\left\{ \frac{X - (1500)(.3)}{\sqrt{(1500)(.3)(.7)}} \geq \frac{500.5 - (1500)(.3)}{\sqrt{(1500)(.3)(.7)}} \right\}$$
$$= P\{Z \geq 2.85\} = 1 - P\{Z \leq 2.85\} = .0022$$