Midterm practice questions

Questions 1-10 are multiple choices, each worth 3 points; Questions 11-17 are free response, each worth 10 points; Total 100 points.

1. The **standard deviation** of a numerical data set measures the ____________ of the data.

a) average

b) most frequent value

c) variability

d) size

e) range

2. The Department of Education wishes to estimate the proportion of all college students who have a job off-campus. It surveyed 1600 randomly selected students; 451 had such jobs. The ***population*** of interest to the Department of Education is:

a) All 1600 students surveyed.

b) The 451 students in the survey who had off-campus jobs.

c) All college students.

d) All college students who have an off-campus job.

e) The Department of Education.

3. Consider the sample data: -3, 1.5, 5, 2.5, 4.5, -1.5. The **mean**, **median** and the **standard deviation** of the data are

a) (-1.5, 4.5, 3.21)

b) (9.0, 2.93, 65)

c) (1.5, 2.25, 2.93)

d) (1.5, 2.0, 3.21)

e) (9.0, 2.0, -3)

4. We have seen that outliers can produce problematic results. Rank the following measures in order or "least affected by outliers" to "most affected by outliers".

a) mean, median, range

b) median, mean, range

c) range, median, mean

d) median, range, mean

e) range, mean, median

5. Consider the result of a STAT exam taken by 120 students, as given in the

following relative frequency distribution:

Grade	Less than 50	50-59	60-69	70-79	80-89	90-100
Cumulative Frequency	15%	10%	30%	25%	15%	5%

How many students received at least a 70 on this exam?

a) 54

b) 45

c) 25

d) 30

e) 66

Questions #6-#8 are based on the following sample of ages (in months) of 18 children at a day care: 36, 42, 18, 32, 22, 22, 25, 29, 30, 31, 19, 24, 35, 29, 26, 36, 24, 28

6. The median age of the children is...

a) 29

b) 28.2

c) 30.5

d) 28.5

e) 31

7. The interquartile range for this data set is...

a) 8

b) 12

c) 16

d) 20

e) 24

8. The standard deviation of the age of children is...

a) 41.24

b) 11.33

c) 10.20

d) 6.42 ✓

e) 6.24

9. If P(A) = 0.7, P(B) = 0.5, and P(A ∩ B) = 0.4, find P(A ∪ B).

a) 0.2

b) 0.3

c) 0.5

d) 0.7

e) 0.8 ✓

10. Two fair, six sided dice are rolled. What is the probability that they show different numbers?

a) 1/2

b) 5/9

c) 2/3

d) 25/36

e) 5/6 ✓

11. Write the definition or formula of

(a) Continuous random variable (3pts)

(b) P(A ∪ B), $P(A \cup B \cup C)$ (3pts)

(c) Bayes formula (4pts) →

Bayes thm.

Let $A_1, A_2, \dots A_k$ be a collection of k mutually exclusive and exhaustive events with prior probability $P(A_i)$ $(i=1,2,\dots k)$. Then for any other event B for which $P(B)>0$. the posterior probability of A_j given that B has occured is

$$P(A_j|B) = \frac{P(A_j \cap B)}{P(B)} = \frac{P(B|A_j)P(A_j)}{\sum_{i=1}^{k} P(B|A_i)P(A_i)} \quad j=1,2,\dots k$$

12. If A and B are independent events, show that A^c and B are also independent. (10pts)

$A \perp B$

$$P(A^c B) = P(B) - P(AB)$$
$$= P(B) - P(B)P(A)$$
$$= P(B)(1-P(A)) = P(B)P(A^c). \blacksquare$$

13. Computer keyboard failures can be attributed to electrical defects or mechanical defects. A repair facility currently has 25 failed keyboards, 6 of which have electrical defects and 19 of which have mechanical defects.

a) In how many ways can a sample of 5 keyboards be selected so that exactly two have an electrical defect? (5pts)

b) If a sample of 5 keyboards is randomly selected, what is the probability that at least 4 of these will have a mechanical defect? (5pts)

a). $\binom{6}{2}\binom{19}{3} = 14535$

b) $p = \frac{\binom{19}{4}\binom{6}{1} + \binom{19}{5}\binom{6}{0}}{\binom{25}{5}} = \frac{34884}{53130} = 0.6566$

14. The random variable X has binomial distribution $Bin(30, 0.3)$, determine the following:

a) $P(X = 11)$ (3pts) 0.1103

b) $P(X < 15)$ (3pts) 0.9831

use Calculator binompdf binomcdf.

c) $P(8 < X \leq 13)$ (4pts)

$F(13) - F(8) = 0.5284$

15.

a) Let $X \sim Bernulli(1,\ p)$, compute $E(X^{79})$? (5pts) $EX^{79} = p$

b) Let $Y \sim Bin(10,\ p)$, compute $E(Y^2)$? (5pts)

$EY^2 = (EY)^2 + V(Y) = (10p)^2 + 10p(1-p)$

$= 100p^2 + 10p - 10p^2 = 90p^2 + 10p$

16. The number of industrial injuries **per working week** in a particular factory is known to follow a Poisson distribution with mean 0.5. Find the probability that

(a) in **a particular week** there will be less than 2 accidents. (5pts)

(b) in a **three-week period,** there will be no accidents. (5pts)

17. A factory produces nails and packs them in boxes of 200. If the probability that a nail is substandard is 0.006, find the probability that a box selected at random contains at most two nails which are substandard. (hint: when n is large and p is small, we can use poisson distribution to approximate binomial distribution) (10 pts)

X: # of substandard nails. $X \sim bin(200, 0.06)$

$P(X \leq 2) = P(X=0) + P(X=1) + P(X=2) = 0.88$

use poisson approximation. X approximately distribute as $pois(200 \times 0.00) = pois(1.2)$

$P(X=0) = e^{-1.2}$ $P(X=1) = e^{-1.2}\frac{1.2}{1!}$ $P(X=2) = e^{-1.2}\frac{1.2}{2!}$

16. Let X be the # of industrial injuries in a week,

then $X \sim pois(0.5)$.

a) $P(X<2) = P(X=0) + P(X=1)$

$= 0.9098$

b) On a particular week, $X \sim pois(0.5)$, then on a 3 week period, Let Y be the # of injuries on the 2nd week and Z be the # of injuries on 3rd week. X, Y, Z are independent. $X, Y, Z \sim pois(0.5)$

$P(\text{no injuries on 3 week periods}) = P(X=0, Y=0, Z=0)$

$= P(X=0)\, P(Y=0)\, P(Z=0)$

$= (e^{-0.5})^3 = 0.2231$