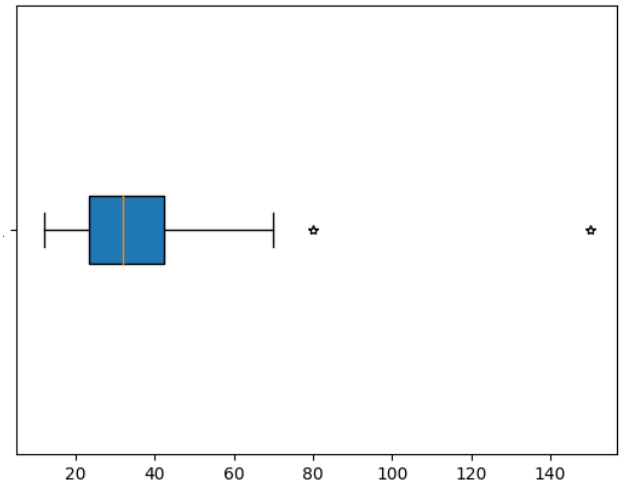
Practice Exam(2)

Q1 Suppose the figure below represent a box plot for following data set:

12, 15, 17, 20, 22, 25, 28, 29, 30, 32, 33, 35, 38, 40, 45, 60, 70, 80, 150

With $Q_1 = 22$, $Q_3 = 45$ and IQR = 23, which of the following is a true statement regard the outliers:

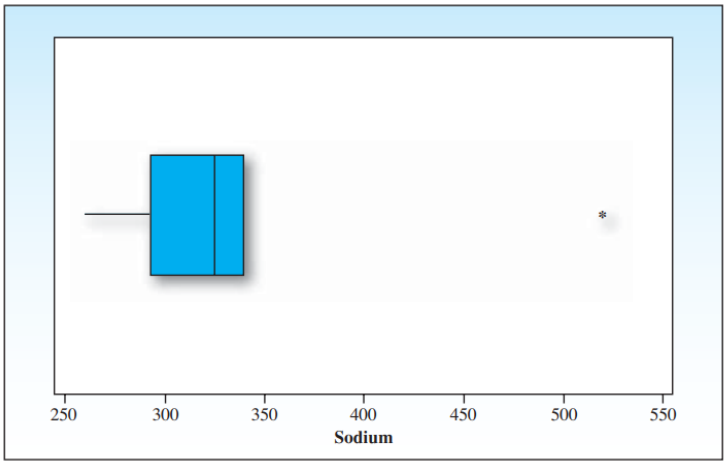
A) Above the upper fence79.5
B) Above the upper fence 85
C) Below 12
D) Below the first quartile
E) Below the second quartile.



Q2 we have the following: 340, 300, 520, 340, 320, 290, 260, 330.

A box plot was constructed for this data, number of outliers is:

A) 3 B) 5 C) 0 D) 1 E) None



Q3 Suppose that $Q_1 = 20$ and $Q_3 = 28$ for a data set. In this case the outliers are the data entries that are

(A) Less than 9 or greater than 30

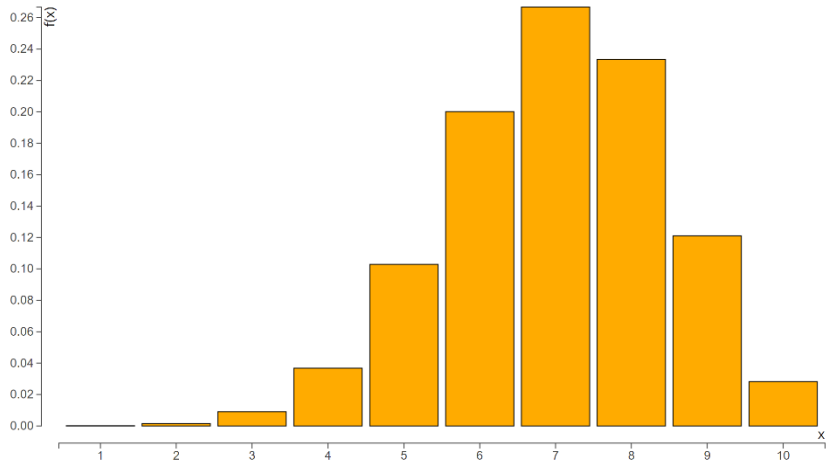
(B) Less than 8 only

(C) Greater than 40 only

(D) Less than 7 or greater than 41

(E) Less than 8 or greater than 40

Q4 suppose the following histogram represents the probability distribution of a Binomial random variable, one of the following represent the values of n and p:

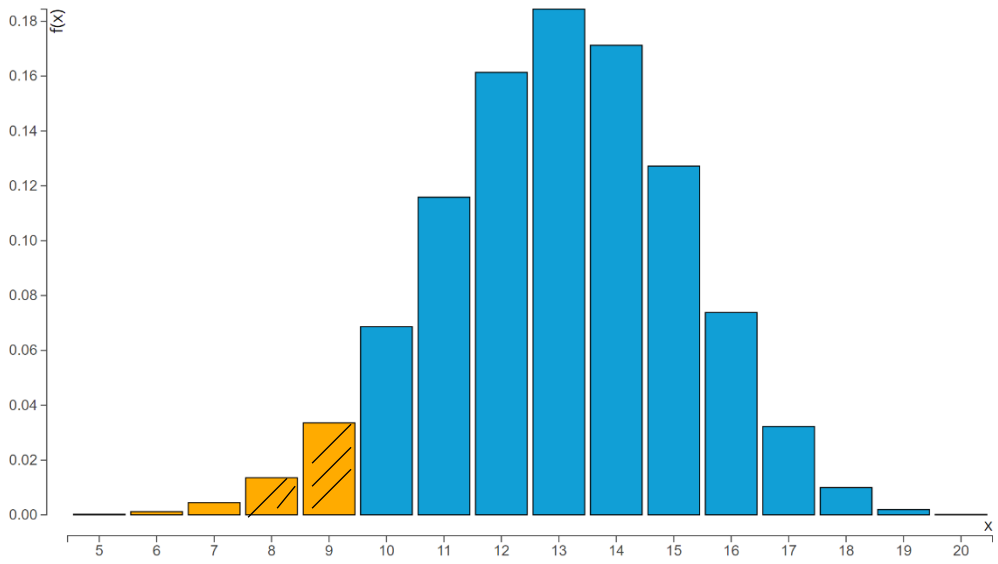


(A) n=10, p = 0.30 (B) n= 15, p=0.7 (C) n = 10, p =0.7 (D) n=10, p=0.95

Q5 Consider a family with a mother, father, and two children. Let A = {mother has influenza}, B = {father has influenza}, C = {first child has influenza}, D= {second child has influenza}, P(A) = P(B) = 0.10, P(A∩ B) = 0.05, then probability the mother has influenza given that the father doesn't:

A) 0.0526 B) 0.001 C) 0.15 D) 0.20 E) 0.0933

Q6 The following is the histogram for Bin(20, 0.65), the shaded bars represent:



A) P(X<4) (B) P(5<X<9) C) P(X≤ 9) D) P(X>16)

Q7 Use the information below.

• The probability that an airplane flight departs on time is 0.89.

• The probability that a flight arrives on time is 0.87.

• The probability that a flight departs and arrives on time is 0.83.

Find the probability that a flight departed on time given that it's not arriving on time.

A) 0.462 B) 0.06 C) 0.13 D) 0.87

Q8 which of the following is a false statement:

I. Adding a constant value to each data point in a dataset will not change the standard deviation.
II. The coefficient of variation is calculated by dividing the standard deviation by the mean.
III. Every independent events are disjoint.
IV. The multiplication rule for independent events states that the probability of both events occurring is the product of their individual probabilities.
V. If events A and B are mutually exclusive, the conditional probability of A given B is always 0.

A) II only B) II and III C) III only D) all are true.

Q9 Suppose that a disease is inherited via a dominant mode of inheritance and that only one of the two parents is affected with the disease. The implications of this mode of inheritance are that the probability is 1 in 2 that any particular offspring will get the disease. If you took two siblings, what is the probability that neither sibling is affected?

A) 0.50 B) 0.75 C) 0.25 D) 0.85 E) 0.30

Q10 Suppose that a disease is inherited via a dominant mode of inheritance and that only one of the two parents is affected with the disease. The implications of this mode of inheritance are that the probability is 1 in 2 that any particular offspring will get the disease. Suppose the older child is affected. What is the probability that the younger child is affected?

A) 0.50 B) 0.25 C) 0.566 D) 0.005

Q11 Suppose a person is regarded as having high cholesterol if $Z > 2.0$ and borderline cholesterol if $1.5 < Z < 2.0$. What proportion of people have high cholesterol?

A) 2.28% B) 97.72% C) 1% D) 0.2280 E) None

Q12 Suppose a person is regarded as having high cholesterol if $Z > 2.0$ and borderline cholesterol if $1.5 < Z < 2.0$. What proportion of people have borderline cholesterol?

A) 97.72% B) 93.32% C) 4.4% D) 2.28%

Q13 Suppose that total carbohydrate intake in 12- to 14-yearold boys is normally distributed, with mean = 124 g/1000 Cal and standard deviation = 20 g/1000 cal. What percentage range have carbohydrate intake above 140 g/1000 cal

A) 0.7881 B) 0.8 C) 0.1829 D) 0.2119

Q14 Serum cholesterol is an important risk factor for coronary disease. We can show that serum cholesterol is approximately normally distributed, with mean = 219 mg/dL and standard deviation = 50 mg/dL. If the clinically desirable range for cholesterol is < 200 mg/dL, What proportion of the general population has borderline high-cholesterol levels—that is, > 200 but < 250 mg/dL?

A) 0.3520 B) 0.6480 C) 0.7324 D) 0.62 E) 0.3520

Q15 In pharmacologic research a variety of clinical chemistry measurements are routinely monitored closely for evidence of side effects of the medication under study. Suppose typical blood-glucose levels are normally distributed, with mean = 90 mg/dL and standard deviation = 38 mg/dL. 1 If the normal range is 65−120 mg/dL, in some studies only values at least 1.5 times as high as the upper limit of normal are identified as abnormal. What percentage of values would fall in this range?

A) 0.9911 B) 0.0089 C) 0.8900 D) 120 E) None

Q16 One interesting phenomenon of bacteriuria is that there is a "turnover"; that is, if bacteriuria is measured on the same woman at two different points in time, the results are not necessarily the same. Assume that 20% of all women who are bacteriuric at time 0 are again bacteriuric at time 1 (1 year later), whereas only 4.2% of women who were not bacteriuric at time 0 are bacteriuric at time 1. Let X be the random variable representing the number of bacteriuric events over the two time periods for 1 woman and still assume that the probability that a woman will be positive for bacteriuria at any one exam is 5%. What is the mean of X?

A) 9.3 B) 1.233 C) 0.4892 D) 0.1

Q17 the following table presents the Prevalence of Alzheimer's disease, we took a married couple, each of whom 75-79 years of age, Suppose the probability that both members of this married couple will have Alzheimer's disease is 0.0015. What is the conditional probability that the man will be affected given that the woman is affected?

Age group	Males	Females
65–69	1.6	0.0
70–74	0.0	2.2
75–79	4.9	2.3
80–84	8.6	7.8
85+	35.0	27.9

A) 0.023 B) 0.0015 C) 0.0652 D) 0.001127

Q18 To study smoking habits in eighth-grade students (ages 12–14), researchers used saliva tests for thiocyanate levels. The students watched a film about detecting recent cigarette use from saliva samples and reported their weekly cigarette consumption. Social pressure often causes underreporting in such studies. Suppose self-reporting is the real result while percent with SCN>100 is the test result. Answer the following questions:

Self-reported cigarettes smoked in past week	Number of students	Percent with SCN ≥ 100 μg/mL
None	1163	3.3
1–4	70	4.3
5–14	30	6.7
15–24	27	29.6
25–44	19	36.8
45+	23	65.2

1) What is the sensitivity of the test for light-smoking students (students who smoke 1-14 cigarettes per week)?
   A) 0.9 B) 0.32 C) 0.50 D) 0.389

2) What is the PV+ of the test?
   A) 0.382 B) 0.48 C) 0.2810 D) 0.139

Question	Answer	Question	Answer
1	A	11	A
2	D	12	C
3	E	13	D
4	C	14	E
5	A	15	B
6	C	16	D
7	A	17	C
8	C	18 (1)	C
9	C	18 (2)	B
10	A		

Answers:

Q1

Lower fence = 22 – 1.5(23) = -12.5 → no outliers.

Upper fence = 45 +1.5(23) = 79.5 → 80 and 150 are outliers. → A

Q2 one outlier only. →D

Q3

Lower fence = 20 – 1.5(8) = 8

Upper fence = 28 +1.5(8) = 40 → E

Q4 (C)

Q5

$\Pr(A\backslash\bar{B}) = \frac{\Pr(A\cap\bar{B})}{\Pr(\bar{B})} = \frac{P(A)-P(A\cap B)}{1-P(B)} = \frac{0.10-0.05}{1-0.05} = 0.0526 \rightarrow A$

Q6 (C)

Q7 P(D) = 0.89, P(A) = 0.87 and P(A∩ D) = 0.83

$P(D\backslash\bar{A}) = \frac{P(D\cap\bar{A})}{P(\bar{A})} = \frac{P(D)-P(A\cap D)}{1-P(A)} = \frac{0.89-0.83}{1-0.87} = 0.462 \rightarrow$A

Q8 (C)

Q9 $P(\bar{A}\cap\bar{B}) = P(\bar{A}) \times P(\bar{B}) = 0.50(0.50) = 0.25 \rightarrow$ C

Q10 $P(A\backslash B) = \frac{P(A\cap B)}{P(B)} = \frac{0.5\times0.5}{0.5} = 0.50 \rightarrow$ A

Q11 P(Z<2.0) = 1- P(Z<2.0) = 1 – 0.9772 = 0.0228 → A

Q12 P(1.5 < Z < 2.0) = P(Z<2.0) – P(Z<1.5) = 0.9772 – 0.9332 = 0.044 = 4.4% → C

Q13 P(X > 140) $P(\frac{x-\mu}{\sigma} > \frac{140-124}{20})$ = P(Z > 0.8)

= 1 – P(Z < 0.8) = 1 – 0.7881 = 0.2119 → D

Q14 P(200 < X < 250) = $P(\frac{200-219}{50} < \frac{x-\mu}{\sigma} < \frac{250-219}{50})$ = P(-0.38 < z < 0.62) =

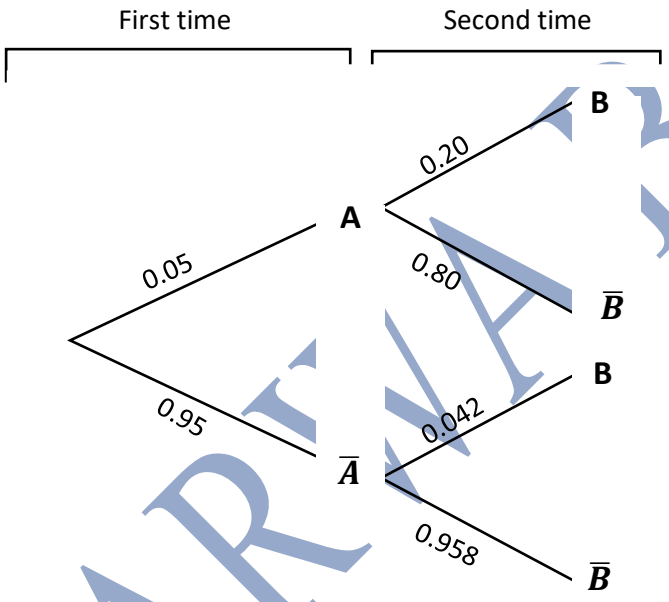
P(Z<0.62) – P(Z < -0.38) = 0.7324 - 0.3520 = 0.3804 → E

Note: P(Z<-0.38) = P(Z > 0.38) = 1 – P(Z< 0.38) = 1 – 0.6480 = 0.3520

Q15 P(X ≥ 1.5(120)) = P(X > 180) = $P(\frac{x-\mu}{\sigma} < \frac{180-90}{38})$ =

P(Z > 2.37) = 1 – P(Z < 2.37) = 1 – 0.9911 = 0.0089 → *B*

Q16



P(X=0) = $P(\bar{A} \cap \bar{B})$ = 0.95(0.958) = 0.9101

P(X=1) = $P(\bar{A} \cap B)$ + $P(A \cap \bar{B})$ = 0.05(0.80) + 0.95 (0.042) = 0.0799 ≈ 0.08

P(X=2) = P(A∩ B) = 0.05 (0.20) = 0.01

X	0	1	2
P(X)	0.9101	0.08	0.01

E(X) = $\sum x.P(X) = 0\ (0.9101) + 1(0.08) + 2(0.01) = 0.1 \rightarrow D$

Q17 P(A) = 0.049 P(B) = 0.023, P(A∩ B) = 0.0015

P(A\B) = $\frac{P(A \cap B)}{P(B)} = \frac{0.0015}{0.023} = 0.0652 \rightarrow C$

Q18

Self-reported	Number of students	Percent	Number of smokers
None	1163	3.3	38.379
1-4	70	4.3	3.01
5-14	30	6.7	2.01
15-24	27	29.6	7.992
25-44	19	36.8	6.992
45+	23	65.2	14.996

1) We can convert our table to the following percentages (we care about light smokers):

   But first 38.379 showed that they smoke but actually they don't so FP = 38.379

   Now the rest of them will be truly negative TN = 1163 – 38.379 = 1124.621

   Now who committed that they smoke, and their test showed they smoke are TP = 3.01 + 2.01 = 5.02

   The rest of them are who smoke but the test showed they don't FN = 100 – 5.02 = 94.98

$$\text{Sensitivity} = \frac{TP}{TP+FN} = \frac{5.02}{5.02+94.98} = 0.0502 \rightarrow C$$

2) Since it's asking for the test

$$\text{P(PV+)} = \frac{TP}{TP+FP} = \frac{3.01+2.01+7.992+6.992+14.996}{38.379+ 3.01+2.01+7.992+6.992+14.996} = 0.48 \rightarrow B$$