IT3030 midterm exam 4/6/2021   ID: _____ Name: _____

1. (25%) The following table describes heart rates for 12 individuals after 10-min running exercise.

patient	Heart rate (beats per minute)
1	131
2	150
3	167
4	150
5	125
6	120
7	150
8	150
9	40
10	136
11	120
12	150

(a) (5%) Determine the mean heart rate.

(b) (5%) Determine the median heart rate.

(c) (10%) Draw a box plot of these 12 observations. Clearly mark Q1, Q2 and Q3, and the two whiskers on the plot.

(d) (5%) Locate the outlier(s) of these observations, if there is any.

[Clearly show the formula/tools/commands you used in getting these results.]

Answer: The sorted list is {40, 120, 120, 125, 131, 136, 150, 150, 150, 150, 150, 167}

(a), (b), (c):

	quartile.exc(0 and 1)	quartile.inc(0 and 1)
average=	132.4166667	
Q1	121.25	123.75
Q2(median)	143	143
Q3	150	150
Box Height	28.75	26.25
min boundary	78.125	84.375
max boundary	193.125	189.375

● Min and max were computed based on 1.5 times the box height.

The smaller whisker would be **120** (the smallest before reaching min)

The larger whisker would be **167** (the largest before reaching max)

(d):

Observation "40" for the 9th individual is an outlier (out of the min-max range).

2. (40%) Among a population of 1,820 people, 30 suffered from tuberculosis (肺結核) and the other 1,790 did not. Chest X-rays were administered to all these people. Among those 30 having the disease, 22 showed positive results and the others negative. Among those healthy individuals, 1,739 showed negative results and the others showed positive.

(a) (10%) Determine TP, FP, TN and FN of this X-rays test.

(b) (10%) Determine the sensitivity, specificity, PPV and NPV of the test.

(c) (10%) What is the probability that a randomly selected individual has tuberculosis that his or her X-ray is positive?

(d) (10%) What is the probability that a randomly selected individual has tuberculosis that his or her X-ray is positive for two consecutive X-ray exams?

Sensitivity = TP/(TP+FN); Specificity = TN/(TN+FP); PPV = TP/(TP+FP); NPV = TN/(TN+FN)

$$P(A_i \mid B) = \frac{P(A_i)P(B \mid A_i)}{P(A_1)P(B \mid A_1) + \ldots + P(A_n)P(B \mid A_n)}$$

Answer:

(a):

TP=22, FN=8, TN=1739, FP=51

(b):

Sensitivity = 22/30 = 0.7333

Specificity = 1739/1790 = 0.9715

PPV=22/(22+51)=22/73=0.3014

NPV=1739/(1739+8)=0.9954

(c):

We need to know P(D+/+)=P(+/D+)P(D+) / [P(+/D+)P(D+) + P(+/D-)P(D-)]

Here P(D+)=30/1820=0.0165, P(D-)=1-0.0165=0.9835

P(+/D+)=sensitivity=0.7333, P(+/D-) = 1-specificity = 0.0285

Therefore P(D+/+) = 0.7333*0.0165 / [0.7333*0.0165 + 0.0285*0.9835]

= 0.0121 / (0.0121 + 0.0280) = 0.0121 / 0.0401 = 0.3017 (This is actually the same as PPV=0.3014, or using MATLAB)

(d):

We need to know P(D+/++)=P(++/D+)P(D+) / [P(++/D+)P(D+) + P(++/D-)P(D-)]

= 0.7333^2*0.0165 / [0.7333^2*0.0165 + 0.0285^2*0.9835]

= 0.0089 / (0.0089 + 0.0008) = 0.0089 / 0.0097 = 0.9175 (or 0.9174 using MATLAB)

3. (20%) In a particular university, 40% of the students are having newspaper reading habit. Nine university students are selected to find their views on reading habit.

(a) Plot the binomial distribution (properly label your x- and y-axis ticks). Determine the probability that:

(b) none of those selected have newspaper reading habit,

(c) all those selected have newspaper reading habit,

(d) at least two third (2/3) of the nine students have newspaper reading habit.

Answer:
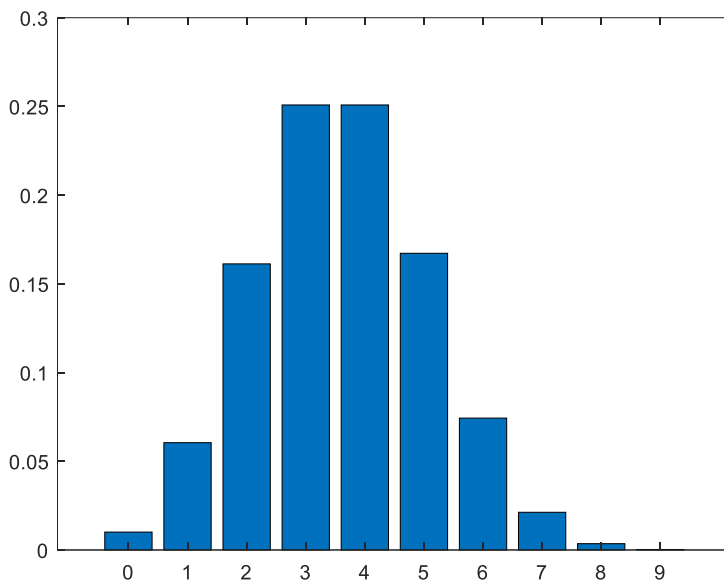
(b) >> binopdf(0,9,0.40) =   0.010077696000000

(c) >> binopdf(9,9,0.40) =   2.621440000000001e-04

(d) >> 1-binocdf(5,9,0.40) =   0.099352576000000

Or:

>> binopdf(6,9,0.4)+binopdf(7,9,0.4)+binopdf(8,9,0.4)+binopdf(9,9,0.4)

  = 0.099352576000000

4. (15%) The patient recovery time from a particular surgical procedure is normally distributed with a mean of 5.3 days and a standard deviation of 2.1 days.

(a) What is the median recovery time?

(b) What is the z-score for a patient who takes ten days to recover?

(c) What is the probability that one patient would need at least 10 days to recover?

Answer:

(a) For normal distribution, the median is the same as the mean, which is 5.3.

(b) >> Z=(10-5.3)/2.1 = 2.2381

(c) >> 1-normcdf(Z) = 0.0126