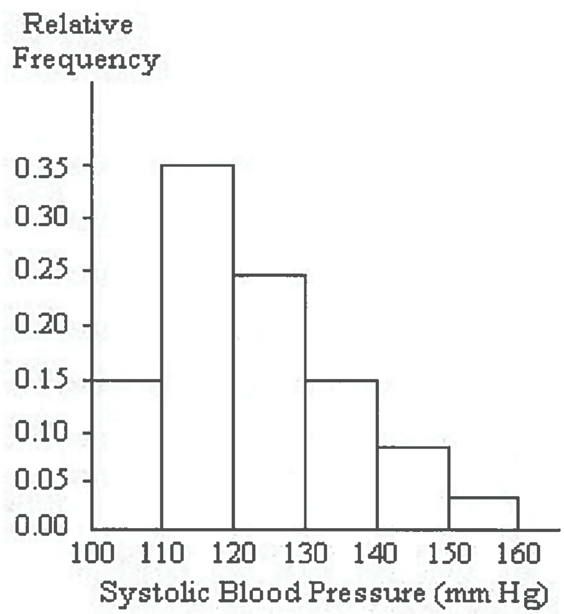
[1] (12 marks) **Circle the best answer.** (2 marks for each question)

Use the following Histogram to answer parts (i) and (iii)

A nurse measured the blood pressure of each person who visited her clinic. Following is a relative-frequency histogram for the systolic blood pressure readings for those people aged between 25 and 40. Use the histogram to answer the question. The blood pressure readings were given to the nearest whole number.



(i) Approximately what percentage of the people aged 25-40 had a systolic blood pressure reading less than 120?

A) 15% B) 75% C) 50% D) 35% E) 25%

$\approx 0.35 + 0.15$
$= 0.5$
$\rightarrow 50\%$

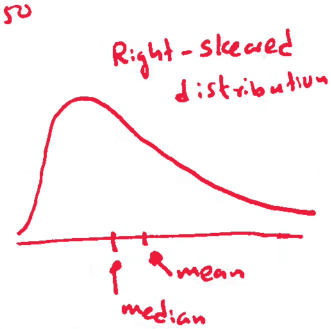
(ii) Given that 200 people were aged between 25 and 40, approximately how many had a systolic blood pressure reading less than 130?

A) 70 B) 25 C) 100 D) 50 E) 150

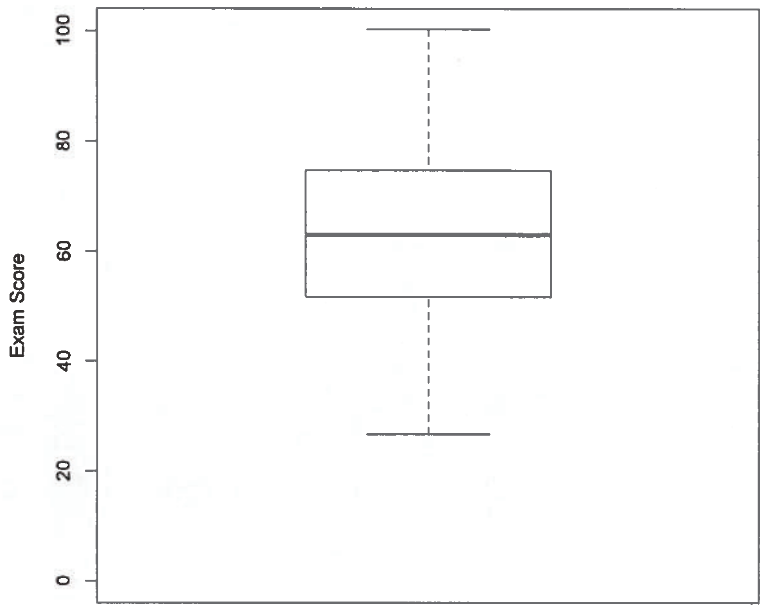
$\approx (0.35+0.25+0.15) \times 200$
$= 0.75 \times 200$
$= 150$

(iii) Which of the following is true for systolic Blood Pressure

A) Mean and Median are same because the histogram is symmetric
B) Mean is greater than median because histogram is left-skewed
C) Mean is greater than median because histogram is right-skewed
D) Median is greater than mean because histogram is left-skewed
E) Median is greater than mean because histogram is right-skewed



The exam scores for all students taking an introductory Statistics course are used to construct the following box plot.



(iv) Based on this box plot, the interquartile range is closest to

A) 10 B) 25 C) 50 D) 80 E) 75

$IQR = Q_3 - Q_1$
$\approx 75 - 50$
$= 25$

(v) If 5 points were added to each score, then

A) the third quartile would increase by 5 points.
B) the median score would increase by 5 points.
C) the interquartile range would remain unchanged.
D) Only A) and B) are correct
E) A), B) and C) all are correct.

(vi) If 5 points were added to each score, then standard deviation of the new scores would
A) be increased by 5.
B) be increased by 25.
C) remain unchanged.
D) be decreased by 5.
E) be decreased by 25.

[2] (7 marks)

(a) If the sample space $S = A \cup B$ and if $P(A) = 0.7$ and $P(B) = 0.6$

i. (3 marks) Find $P(A \cap B)$

Since sample space $S = A \cup B$, $P(A \cup B) = 1$. ← 1

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$ ← 1

$1 = 0.7 + 0.6 - P(A \cap B)$

$P(A \cap B) = 1.3 - 1$

$P(A \cap B) = \underline{\underline{0.3}}$ ← 1

ii. (2 marks) Are A and B independent? Why or why not?

If A and B are independent, $P(A \cap B) = P(A) \cdot P(B)$

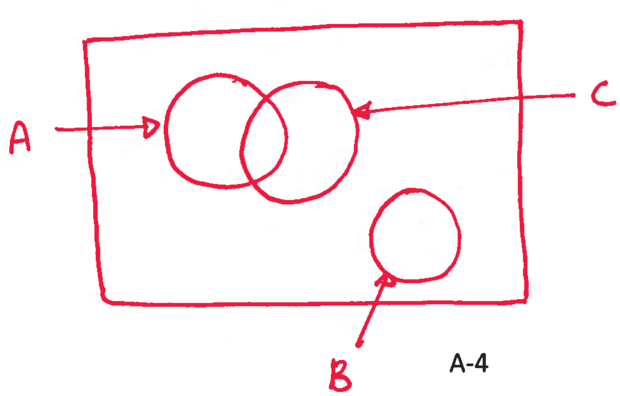
$P(A \cap B) = 0.3$

$P(A) \cdot P(B) = 0.7 \times 0.6 = 0.42$ } 1

$\therefore P(A \cap B) \neq P(A) \cdot P(B)$

Therefore A and B are not independent. } 1

(b) (2 marks) This is not related to Part (a). Draw a Venn Diagram involving sets A, B, and C where A and B are mutually exclusive, B and C are disjoint and $A \cap C \neq \emptyset$



0 or 2 marks.

↑

Correct answers

[3] (7 marks) Let X denote the distance (m) that an animal moves from its birth site to the first territorial vacancy it encounters. Suppose that banner-tailed kangaroo rats, X has an exponential distribution with parameter λ=0.013

(a) (3 marks) What is the probability that the distance is between 100m and 220m

$X \sim exp(\lambda) \Rightarrow f(x) = \lambda e^{-\lambda x} \quad ; x \geq 0 \qquad \lambda = 0.013$

$$P(100 < X < 220) = \int_{100}^{220} \lambda e^{-\lambda x} dx = \frac{\lambda e^{-\lambda x}}{-\lambda}\Big|_{100}^{220} = -e^{-\lambda x}\Big|_{100}^{220} \qquad \leftarrow 1$$

$$= -e^{-220\lambda} - (-e^{-100\lambda})$$

$$= e^{-100\lambda} - e^{-220\lambda} \quad ; \lambda = 0.013 \qquad \leftarrow 1$$

$$= e^{-1.3} - e^{-2.86}$$

$$= 0.2725 - 0.0573$$

$$= 0.2152 \qquad \leftarrow 1$$

(b) (3 marks) What is the value of the median distance?

$$F(x) = P(X \leq x) = \int_0^x \lambda e^{-\lambda t} dt = \frac{\lambda e^{-\lambda t}}{-\lambda}\Big|_0^x = -e^{-\lambda t}\Big|_0^x$$

$$= -e^{-\lambda x} - (-e^0) = 1 - e^{-\lambda x}.$$

Let median is m. Then

$F(m) = 0.5$ 1

$1 - e^{-\lambda m} = 0.5 \Rightarrow -e^{-\lambda m} = -0.5 \Rightarrow -\lambda m = \ln(0.5)$ 1

$m = \frac{\ln(0.5)}{-\lambda} = \frac{\ln(0.5)}{-0.013}$

$= 53.31$ m 1

(c) (1 mark) In language that is clear and accurate, provide an interpretation of the probability expression $P(X < 150|X > 100) = 0.478$

When it is given that a banner-tailed kangaroo rat moves more than 100 m from its birth site to the first territorial vacancy, the probability it moves less than 150 m is 0.478 1

[4] (6 marks) Suppose that $X \sim N(6.3\,,4)$. Using the section of the standard normal table below, calculate $P(\,|X-5| > 2)$.

$X \sim N(6.3, 4)$

$\Rightarrow X \sim N(6.3, 2^2)$

$\mu = 6.3$

$\sigma = 2$

Table entry for z is the area under the standard normal curve to the left of z.

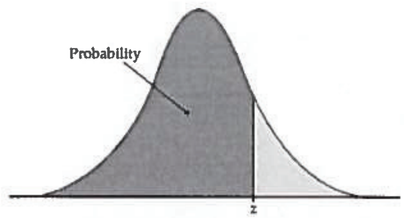


TABLE A

Standard normal probabilities (continued)

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767

$P(|X-5|>2) = 1 - P(|X-5|<2)$

$= 1 - P(-2 < X-5 < 2)$ ← dealing with absolute values 1

$= 1 - P(3 < X < 7)$

$= 1 - P\left(\frac{3-6.3}{2} < \frac{X-6.3}{2} < \frac{7-6.3}{2}\right)$ ← figurin out $\sigma = 2$ 1

$= 1 - P(-1.65 < Z < 0.35)$ ← z values 1

$= 1 - [P(Z<0.35) - P(Z<-1.65)]$

$= 1 - [P(Z<0.35) - P(Z>1.65)]$ ← using symmetry 1

$= 1 - \{P(Z<0.35) - [1 - P(Z<1.65)]\}$

$= 1 - \{0.6368 - 1 + 0.9505]$ ← Correct table values 1

$= 0.4127$ 1 for the Answer

[5] (8 marks) Suppose X_1 and X_2 are two independent random variables, where $X_1 \sim U(10, 25)$ and $X_2 \sim U(20, 30)$. Let $Y = max(X_1, X_2)$. Showing all required steps, find the *pdf* of Y.

$X_1 \sim U(10,25)$

$f_{X_1}(x_1) = \frac{1}{25-10} = \frac{1}{15}$; $10 \le x_1 \le 15$

$X_2 \sim U(20,30)$

$f_{X_2}(x_2) = \frac{1}{30-20} = \frac{1}{10}$; $20 \le x_2 \le 30$

Integrating pdf of X_1 & X_2, we get cdf of X_1 & X_2

$F_{X_1}(x_1) = \int_{10}^{x_1} \frac{1}{15}\, dt = \frac{1}{15} t \Big|_{10}^{x_1} = \frac{x_1 - 10}{15}$ 1 ; $10 \le x_1 \le 25$

similarly, $F_{X_2}(x_2) = \frac{x_2 - 20}{10}$; $20 \le x_2 \le 30$

1

$Y = max(X_1, X_2)$

cdf of Y is $F_Y(y) = P(Y \le y) = P(X_1 \le y, X_2 \le y)$ 1

$= P(X_1 \le y)\, P(X_2 \le y)$ 1 ← because X_1 & X_2 are independent

$= F_{X_1}(y)\, F_{X_2}(y)$

$= \left(\frac{y-10}{15}\right)\left(\frac{y-20}{10}\right)$ 1 ; $20 \le y \le 25$

$$F_Y(y) = \begin{cases} 0 & ; \ y < 20 \\ \left(\frac{y-10}{15}\right)\left(\frac{y-20}{10}\right) & ; \ 20 \le y < 25 \\ \left(\frac{y-20}{10}\right) & ; \ 25 \le y < 30 \\ 1 & ; \ y > 30 \end{cases}$$

1

1 for correct support

pdf of Y is

$f_Y(y) = \frac{d}{dy}\left(F_Y(y)\right)$

$$f_Y(y) = \begin{cases} 0 & ; \ y < 20 \\ \frac{2y-30}{150} & ; \ 20 \le y < 25 \\ \frac{1}{10} & ; \ 25 \le y < 30 \\ 0 & ; \ y > 30 \end{cases}$$

$$\Rightarrow f_Y(y) = \begin{cases} \frac{2y-30}{150} & ; \ 20 \le y \le 25 \\ \frac{1}{10} & ; \ 25 \le y < 30 \\ 0 & ; \ \text{otherwise} \end{cases}$$

1