University of Victoria - Stat 260 - Spring 2023
Term Test 1 - Version A
Section A01 - Instructor: Dr. Michelle Edwards

Instructions:

- This paper has 6 pages of questions plus this cover page. Please count your pages and report any discrepancy immediately to the invigilator.
- Answer all questions on the test paper in the space provided. Questions 1 through 8 are short answer and worth 3 marks each. For each question show your work and write your final answer in the space as indicated. Question 9 is a full-answer question. To obtain full marks on this question, clear step-by-step reasoning must be displayed. Marks will be deducted for incoherent or incomplete solutions.
- Unless an exact answer with fewer decimals can be found, give all answers with at least three decimals of accuracy.
- A Sharp EL-510R series calculator may be used. No other calculator may be used.
- A formula sheet is provided. No other aids such as books, notes, scrap paper, electronic devices, etc. are permitted.
- You have 50 minutes to write the test.

Formula List

$$\frac{\displaystyle\sum_{i=1}^{n}(x_i-\overline{x})^2}{n-1}=\frac{\left(\displaystyle\sum_{i=1}^{n}x_i^2\right)-n\left(\overline{x}\right)^2}{n-1}$$

$$E(X^2)-\mu^2$$

$$\frac{\sum_{i=1}^{n}[(x_i-\bar{x})(y_i-\bar{y})]}{\sqrt{[\sum_{i=1}^{n}(x_i-\bar{x})^2][\sum_{i=1}^{n}(y_i-\bar{y})^2]}}$$

$$\frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}}$$

$$\binom{n}{x}p^x(1-p)^{n-x}, x=0,1,2,\ldots,n$$

$$\frac{\lambda^x}{x!}e^{-\lambda}, x=0,1,\ldots$$

$$1-e^{-\lambda x}, x>0$$

$$E(XY)-\mu_X\mu_Y$$

$$P\left(Z\leq\tfrac{x+0.5-np}{\sqrt{np(1-p)}}\right)$$

$$\frac{x^{\alpha-1}e^{-x/\beta}}{\beta^\alpha\Gamma(\alpha)}$$

$$\textstyle\sum_{i=1}^{n}a_i^2\mathrm{Var}(X_i)+2\sum_{i<j}a_ia_j\mathrm{Cov}(X_i,X_j)$$

$$\frac{s}{\sqrt{n}}$$

$$\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{m}+\frac{\hat{p}_2(1-\hat{p}_2)}{n}}$$

$$\sqrt{\frac{s_1^2}{m}+\frac{s_2^2}{n}}$$

$$\sqrt{\frac{(m-1)s_1^2+(n-1)s_2^2}{m+n-2}\left(\frac{1}{m}+\frac{1}{n}\right)}$$

integer part of $\dfrac{(s_1^2/m+s_2^2/n)^2}{\dfrac{(s_1^2/m)^2}{m-1}+\dfrac{(s_2^2/n)^2}{n-1}}$

estimate $\pm$ (c.v.)(e.s.e.)

$$\frac{\text{estimate} - \text{param. value under } H_0}{\text{e.s.e. or (s.e. under } H_0)}$$

[3] **1.** A study of noise levels at Skytrain stations in Vancouver was performed where the maximum noise level per day (measured in dB) was recorded. The data for a sample of 10 Skytrain stations are as follows:

82 79 74 68 67 81 90 92 73 81

Fill in each blank with the statistic asked for. It is only necessary to show work for part (ii).

(i) The sample mean, $\overline{x}$. 78.7

(ii) The sample median, $\widetilde{x}$. 80

(iii) The sample standard deviation, s. 8.3805

sorted: 67, 68, 73, 74, 79, 81, 81, 82, 90, 92

$$\tilde{x} = \frac{79+81}{2} = 80$$

[3] **2.** Use the blank to indicate whether each statement is true or false. No reasons are necessary.

false (i) The covariance s_{xy} can indicate the strength of a linear relationship between x and y.

false (ii) The correlation coefficient r is equal to the slope of the straight line which best fits a sample of bivariate data.

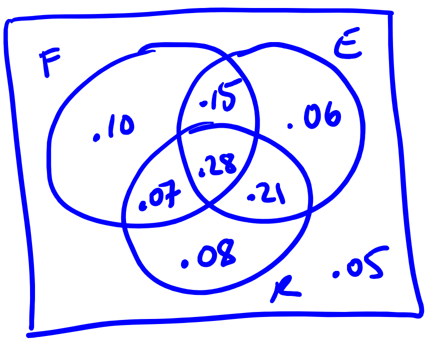
true (iii) A correlation coefficient of $r = 0$ indicates that there is no linear relationship between a sample of bivariate data.

[3] **3.** A ferry boat service keeps statistics on three characteristics: the type of vehicles customers drive, if they arrive at least 45 minutes in advance of their sailing, and if they make a reservation when sailing. It is determined that:

F = four-door vehicle
E = at least 45 minutes in advance
R = reservation

- 60% drive a four-door vehicle
- 70% arrive at least 45 minutes in advance
- 64% make a reservation
- 43% drive a four-door vehicle and arrive at least 45 minutes in advance
- 35% drive a four-door vehicle and make a reservation
- 49% arrive at least 45 minutes in advance and make a reservation
- 28% drive a four-door vehicle, arrive at least 45 minutes in advance, and make a reservation

Suppose a customer travelling on the ferry boat is selected at random. What is the probability they have exactly two of the characteristics? (For example, one such case is that the customer drives a four-door vehicle and arrives at least 45 minutes in advance of their sailing, but they did not make a reservation.)



$P(\text{exactly two})$
$= P(F \cap E \cap \bar{R}) + P(F \cap \bar{E} \cap R) + P(\bar{F} \cap E \cap R)$
$= .15 + .07 + .21$
$= .43$

Answer:

0.43

[3] **4.** A car mechanic has noticed that many cars being serviced have tires that are under-filled with air. Suppose the probability that a car has an under-filled tire is 54%.
The mechanic is scheduled to service five cars in one day. Assuming that cars have under-filled tires independent of one another, what is the probability that at least one of the five cars will have an under-filled tire?

$P(\text{at least one underfilled})$
$= 1 - P(\text{none underfilled})$
$= 1 - (1 - .54)^5$
$= 1 - (0.46)^5$
$= 0.9794$

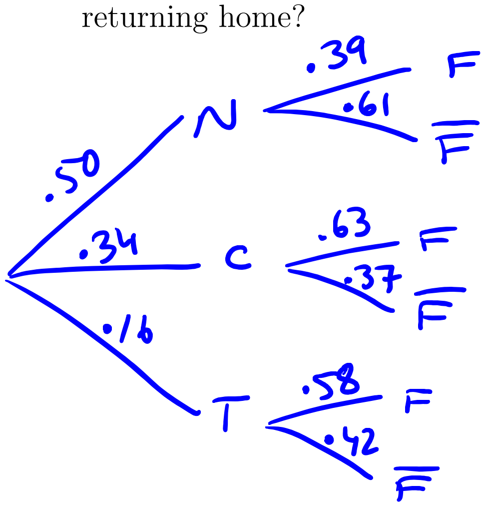
Answer:

0.9794

(Questions 5 and 6 refer to the following setup)
A pilot from Victoria is planning an upcoming flight. For all the flights they have logged on Vancouver Island, 50% have been to Nanaimo, 34% have been to Comox, and the rest have been to Tofino.

Of the flights that are to Nanaimo, for 39% of them the pilot refuels the plane before returning home. Of the flights that are to Comox, for 63% of them the pilot refuels the plan before returning home. Of the flights that are to Tofino, for 58% of them the pilot refuels the plane before returning home.

[3] **5.** For their upcoming flight, what is the probability that the pilot does not refuel the plane before returning home?



Answer:

0.498

$$P(\bar{F}) = (.50)(.61) + (.34)(.37) + (.16)(.42)$$
$$= 0.498$$

[3] **6.** Suppose an upcoming flight is one where the pilot refuels the plane before returning home. What is the probability the flight is to Comox?

$$P(C \mid F) = \frac{P(C \cap F)}{P(F)}$$

$$= \frac{(.34)(.63)}{(.50)(.39) + (.34)(.63) + (.16)(.58)}$$

$$= \frac{0.2142}{0.502} = 0.4267$$

Answer:

0.4267

(Questions 7 and 8 refer to the following setup)
A large factory estimates that the number of employees which stay home sick on a given day has the following cumulative distribution function:

	x	0	1	2	3	4	5	6	7
$P(X\le x)=$	$F(x)$	0.03	0.08	0.19	0.38	0.54	0.76	0.91	1
$P(X=x)=$	$f(x)$	.03	.05	.11	.19	.16	.22	.15	.09

[3] **7.** If it is known that on a random day at least 2 employees stays home sick, what is the probability that no more than 5 employees stay home sick?

Answer: 0.7391

$$P(X\le 5 \mid X\ge 2) = \frac{P(2\le X\le 5)}{P(X\ge 2)}$$

$$= \frac{P(X\le 5)-P(X\le 1)}{1-P(X\le 1)} = \frac{0.76-0.08}{1-0.08} = \frac{0.68}{0.92} = 0.7391$$

or $\frac{P(2\le X\le 5)}{P(X\ge 2)} = \frac{0.11+0.19+0.16+0.22}{0.11+0.19+0.16+0.22+0.15+0.09} = \frac{0.68}{0.92} = 0.7391$

[3] **8.** What is the expected number of employees that stay home sick on a given day?

Answer: 4.11

$$E(X) = \sum x\cdot f(x)$$

$$= 0(.03)+1(.05)+2(.11)+3(.19)$$

$$+4(.16)+5(.22)+6(.15)+7(.09)$$

$$= 4.11$$

(Question 9) The mountain pine beetle is a species of beetle native to forests in western North America. When the pine beetles lay their eggs under the tree bark they introduce blue stain fungus which blocks water and nutrient transport within the tree, causing it to die. In a specific area of forest the trees are assessed for damage and are categorized as 0 - no damage, 1 - moderate damage, or 2 - severe damage. The trees in the area will be used immediately for lumber, for industrial use, or left for future growth. The proportion of each classification is summarized below:

L = lumber
I = industrial
F = future

		Amount of Damage			
		0	1	2	
Use	lumber	0.10	0.12	0.15	0.37
	industrial	0.23	0.11	0.09	0.43
	future growth	0.08	0.07	0.05	0.20
		0.41	0.30	0.29	

(a) **[2 marks]** What is the probability that a randomly selected tree from this area has moderate damage and will not be left for future growth?

$P(1 \cap F) = 0.12 + 0.11 = 0.23$

(b) **[2 marks]** Are the events "tree has moderate damage" and "tree is left for future growth" independent? Support your answer with a calculation.

$P(1 \cap F) = 0.07$

$P(1) \cdot P(F) = (0.30)(0.20) = 0.06$

$P(1 \cap F) \neq P(1) \cdot P(F)$ so the events are not independent.

The table from the previous page is copied below. (Recall that 0 - no damage, 1 - moderate damage, and 2 - severe damage.)

		Amount of Damage			
		0	1	2	
Use	lumber	0.10	0.12	0.15	0.37
	industrial	0.23	0.11	0.09	0.43
	future growth	0.08	0.07	0.05	0.20
		0.41	0.30	0.29	

(c) **[2 marks]** If a randomly selected tree is found to have no damage, what is the probability that it is used for lumber or for industrial use?

$$P(L \cup I \mid 0) = \frac{P((L \cup I) \cap 0)}{P(0)} = \frac{0.10 + 0.23}{0.10 + 0.23 + 0.08}$$

$$= \frac{0.33}{0.41} = 0.8049$$

University of Victoria - Stat 260 - Spring 2023
Term Test 1 - Version B
Section A01 - Instructor: Dr. Michelle Edwards

Instructions:

- This paper has 6 pages of questions plus this cover page. Please count your pages and report any discrepancy immediately to the invigilator.
- Answer all questions on the test paper in the space provided. Questions 1 through 8 are short answer and worth 3 marks each. For each question show your work and write your final answer in the space as indicated. Question 9 is a full-answer question. To obtain full marks on this question, clear step-by-step reasoning must be displayed. Marks will be deducted for incoherent or incomplete solutions.
- Unless an exact answer with fewer decimals can be found, give all answers with at least three decimals of accuracy.
- A Sharp EL-510R series calculator may be used. No other calculator may be used.
- A formula sheet is provided. No other aids such as books, notes, scrap paper, electronic devices, etc. are permitted.
- You have 50 minutes to write the test.

Formula List

$$\frac{\sum_{i=1}^{n}(x_i-\overline{x})^2}{n-1}=\frac{\left(\sum_{i=1}^{n}x_i^2\right)-n\left(\overline{x}\right)^2}{n-1}$$

$$E(X^2)-\mu^2$$

$$\frac{\sum_{i=1}^{n}[(x_i-\bar{x})(y_i-\bar{y})]}{\sqrt{[\sum_{i=1}^{n}(x_i-\bar{x})^2][\sum_{i=1}^{n}(y_i-\bar{y})^2]}}$$

$$\frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}}$$

$$\binom{n}{x}p^x(1-p)^{n-x}, x=0,1,2,\ldots,n$$

$$\frac{\lambda^x}{x!}e^{-\lambda}, x=0,1,\ldots$$

$$1-e^{-\lambda x}, x>0$$

$$E(XY)-\mu_X\mu_Y$$

$$P\left(Z\le\frac{x+0.5-np}{\sqrt{np(1-p)}}\right)$$

$$\frac{x^{\alpha-1}e^{-x/\beta}}{\beta^\alpha\Gamma(\alpha)}$$

$$\sum_{i=1}^{n}a_i^2\mathrm{Var}(X_i)+2\sum_{i<j}a_ia_j\mathrm{Cov}(X_i,X_j)$$

$$\frac{s}{\sqrt{n}}$$

$$\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{m}+\frac{\hat{p}_2(1-\hat{p}_2)}{n}}$$

$$\sqrt{\frac{s_1^2}{m}+\frac{s_2^2}{n}}$$

$$\sqrt{\frac{(m-1)s_1^2+(n-1)s_2^2}{m+n-2}\left(\frac{1}{m}+\frac{1}{n}\right)}$$

$$\text{integer part of }\frac{(s_1^2/m+s_2^2/n)^2}{\frac{(s_1^2/m)^2}{m-1}+\frac{(s_2^2/n)^2}{n-1}}$$

estimate $\pm$ (c.v.)(e.s.e.)

$$\frac{\text{estimate}-\text{param. value under }H_0}{\text{e.s.e. or (s.e. under }H_0)}$$

[3] **1.** A study of noise levels at Skytrain stations in Vancouver was performed where the maximum noise level per day (measured in dB) was recorded. The data for a sample of 10 Skytrain stations are as follows:

84 ~~76~~ ~~70~~ ~~65~~ ~~68~~ 83 91 94 ~~70~~ ~~74~~

Fill in each blank with the statistic asked for. It is only necessary to show work for part (ii).

(i) The sample mean, $\overline{x}$. 77.5

(ii) The sample median, $\widetilde{x}$. 75

(iii) The sample standard deviation, s. 10.0028

sorted: 65, 68, 70, 70, 74, 76, 83, 84, 91, 94

$$\tilde{x} = \frac{74+76}{2} = 75$$

[3] **2.** Use the blank to indicate whether each statement is true or false. No reasons are necessary.

true (i) The covariance s_{xy} can indicate the direction of a linear relationship between x and y. (i.e. positive or negative)

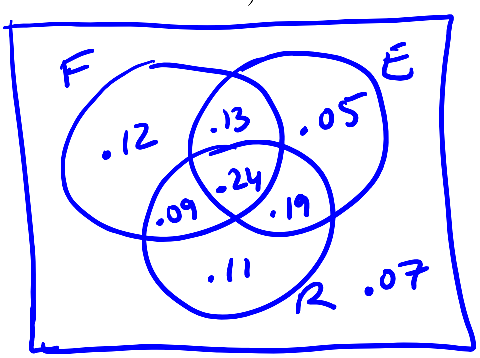
false (ii) A correlation coefficient of $r = 0$ indicates that there is no relationship between a sample of bivariate data.

false (iii) The correlation coefficient r is equal to the slope of the straight line which best fits a sample of bivariate data.

[3] **3.** A ferry boat service keeps statistics on three characteristics: the type of vehicles customers drive, if they arrive at least 45 minutes in advance of their sailing, and if they make a reservation when sailing. It is determined that:

- 58% drive a four-door vehicle — F = four-door vehicle
- 61% arrive at least 45 minutes in advance — E = at least 45 minutes in advance
- 63% make a reservation — R = reservation
- 37% drive a four-door vehicle and arrive at least 45 minutes in advance
- 33% drive a four-door vehicle and make a reservation
- 43% arrive at least 45 minutes in advance and make a reservation
- 24% drive a four-door vehicle, arrive at least 45 minutes in advance, and make a reservation

Suppose a customer travelling on the ferry boat is selected at random. What is the probability they have exactly two of the characteristics? (For example, one such case is that the customer drives a four-door vehicle and arrives at least 45 minutes in advance of their sailing, but they did not make a reservation.)



$P(\text{exactly two})$
$= P(F \cap E \cap \bar{R}) + P(F \cap \bar{E} \cap R) + P(\bar{F} \cap E \cap R)$
$= 0.13 + 0.09 + 0.19$
$= 0.41$

Answer: 0.41

[3] **4.** A car mechanic has noticed that many cars being serviced have tires that are under-filled with air. Suppose the probability that a car has an under-filled tire is 47%.
The mechanic is scheduled to service five cars in one day. Assuming that cars have under-filled tires independent of one another, what is the probability that at least one of the five cars will have an under-filled tire?

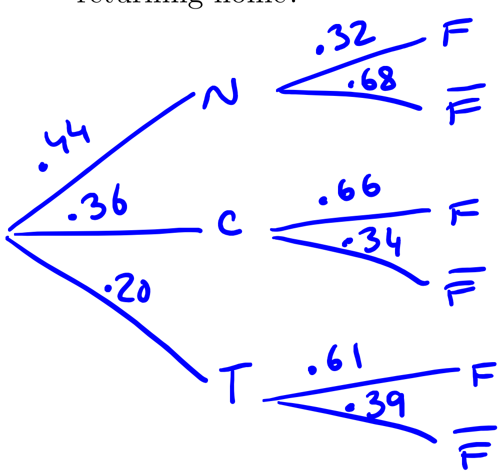
$P(\text{at least one underfilled})$
$= 1 - P(\text{none underfilled})$
$= 1 - (1-0.47)^5$
$= 1 - (0.53)^5$
$= 0.9582$

Answer: 0.9582

(Questions 5 and 6 refer to the following setup)
A pilot from Victoria is planning an upcoming flight. For all the flights they have logged on Vancouver Island, 44% have been to Nanaimo, 36% have been to Comox, and the rest have been to Tofino.

Of the flights that are to Nanaimo, for 32% of them the pilot refuels the plane before returning home. Of the flights that are to Comox, for 66% of them the pilot refuels the plan before returning home. Of the flights that are to Tofino, for 61% of them the pilot refuels the plane before returning home.

[3] **5.** For their upcoming flight, what is the probability that the pilot does not refuel the plane before returning home?



Answer

0.4996

$$P(\bar{F}) = (.44)(.68) + (.36)(.34) + (.20)(.39)$$
$$= 0.4996$$

[3] **6.** Suppose an upcoming flight is one where the pilot refuels the plane before returning home. What is the probability the flight is to Tofino?

$$P(T \mid F) = \frac{P(T \cap F)}{P(F)}$$

$$= \frac{(.20)(.61)}{(.44)(.32) + (.36)(.66) + (.20)(.61)}$$

$$= \frac{0.122}{0.5004} = 0.2438$$

Answer

0.2438

(Questions 7 and 8 refer to the following setup)
A large factory estimates that the number of employees which stay home sick on a given day has the following cumulative distribution function:

	x	0	1	2	3	4	5	6	7
$P(X \le x) =$	$F(x)$	0.04	0.09	0.24	0.35	0.53	0.75	0.88	1
$P(X = x) =$	$f(x)$	.04	.05	.15	.11	.18	.22	.13	.12

[3] **7.** If it is known that on a random day at least 3 employees stays home sick, what is the probability that no more than 6 employees stay home sick?

$$P(X \le 6 \mid X \ge 3) = \frac{P(3 \le X \le 6)}{P(X \ge 3)}$$

$$= \frac{P(X \le 6) - P(X \le 2)}{1 - P(X \le 2)} = \frac{0.88 - 0.24}{1 - 0.24} = \frac{0.64}{0.76} = 0.8421$$

or $$\frac{P(3 \le X \le 6)}{P(X \ge 3)} = \frac{0.11 + 0.18 + 0.22 + 0.13}{0.11 + 0.18 + 0.22 + 0.13 + 0.12} = \frac{0.64}{0.76} = 0.8421$$

Answer
0.8421

[3] **8.** What is the expected number of employees that stay home sick on a given day?

$$E(X) = \sum x \cdot f(x)$$

$$= 0(.04) + 1(.05) + 2(.15) + 3(.11)$$

$$+ 4(.18) + 5(.22) + 6(.13) + 7(.12)$$

$$= 4.12$$

Answer
4.12

(Question 9) The mountain pine beetle is a species of beetle native to forests in western North America. When the pine beetles lay their eggs under the tree bark they introduce blue stain fungus which blocks water and nutrient transport within the tree, causing it to die. In a specific area of forest the trees are assessed for damage and are categorized as 0 - no damage, 1 - moderate damage, or 2 - severe damage. The trees in the area will be used immediately for lumber, for industrial use, or left for future growth. The proportion of each classification is summarized below:

L = lumber
I = industrial
F = future growth

		Amount of Damage			
		0	1	2	
Use	lumber	0.13	0.11	0.06	0.30
	industrial	0.20	0.12	0.09	0.41
	future growth	0.10	0.14	0.05	0.29
		0.43	0.37	0.20	

(a) **[2 marks]** What is the probability that a randomly selected tree from this area has no damage and will not be used for industrial use?

$P(0 \cap \bar{I}) = 0.13 + 0.10 = 0.23$

(b) **[2 marks]** Are the events "tree has severe damage" and "tree is used for lumber" independent? Support your answer with a calculation.

$P(2 \cap L) = 0.06$

$P(2) \cdot P(L) = (0.20)(0.30) = 0.06$

$P(2 \cap L) = P(2) \cdot P(L)$ so the events <u>are</u> independent.

The table from the previous page is copied below. (Recall that 0 - no damage, 1 - moderate damage, and 2 - severe damage.)

		Amount of Damage			
		0	1	2	
Use	lumber	0.13	0.11	0.06	0.30
	industrial	0.20	0.12	0.09	0.41
	future growth	0.10	0.14	0.05	0.29
		0.43	0.37	0.20	

(c) **[2 marks]** If a randomly selected tree is found to have moderate damage, what is the probability that it is used for industrial use or is left for future growth?

$$P(I \cup F \mid 1) = \frac{P((I \cup F) \cap 1)}{P(1)} = \frac{0.12 + 0.14}{0.11 + 0.12 + 0.14}$$

$$= \frac{0.26}{0.37} = 0.7027$$

University of Victoria - Stat 260 - Spring 2023
Term Test 1 - Version A
Section A02 - Instructor: Dr. Michelle Edwards

Instructions:

- This paper has 6 pages of questions plus this cover page. Please count your pages and report any discrepancy immediately to the invigilator.
- Answer all questions on the test paper in the space provided. Questions 1 through 8 are short answer and worth 3 marks each. For each question show your work and write your final answer in the space as indicated. Question 9 is a full-answer question. To obtain full marks on this question, clear step-by-step reasoning must be displayed. Marks will be deducted for incoherent or incomplete solutions.
- Unless an exact answer with fewer decimals can be found, give all answers with at least three decimals of accuracy.
- A Sharp EL-510R series calculator may be used. No other calculator may be used.
- A formula sheet is provided. No other aids such as books, notes, scrap paper, electronic devices, etc. are permitted.
- You have 50 minutes to write the test.

Formula List

$$\frac{\sum_{i=1}^{n}(x_i-\overline{x})^2}{n-1}=\frac{\left(\sum_{i=1}^{n}x_i^2\right)-n\left(\overline{x}\right)^2}{n-1}$$

$$E(X^2)-\mu^2$$

$$\frac{\sum_{i=1}^{n}[(x_i-\bar{x})(y_i-\bar{y})]}{\sqrt{[\sum_{i=1}^{n}(x_i-\bar{x})^2][\sum_{i=1}^{n}(y_i-\bar{y})^2]}}$$

$$\frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}}$$

$$\binom{n}{x}p^x(1-p)^{n-x}, x=0,1,2,\ldots,n$$

$$\frac{\lambda^x}{x!}e^{-\lambda}, x=0,1,\ldots$$

$$1-e^{-\lambda x}, x>0$$

$$E(XY)-\mu_X\mu_Y$$

$$P\left(Z\leq \tfrac{x+0.5-np}{\sqrt{np(1-p)}}\right)$$

$$\frac{x^{\alpha-1}e^{-x/\beta}}{\beta^{\alpha}\Gamma(\alpha)}$$

$$\textstyle\sum_{i=1}^{n}a_i^2\mathrm{Var}(X_i)+2\sum_{i<j}a_ia_j\mathrm{Cov}(X_i,X_j)$$

$$\frac{s}{\sqrt{n}}$$

$$\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{m}+\frac{\hat{p}_2(1-\hat{p}_2)}{n}}$$

$$\sqrt{\frac{s_1^2}{m}+\frac{s_2^2}{n}}$$

$$\sqrt{\frac{(m-1)s_1^2+(n-1)s_2^2}{m+n-2}\left(\frac{1}{m}+\frac{1}{n}\right)}$$

$$\text{integer part of }\frac{(s_1^2/m+s_2^2/n)^2}{\frac{(s_1^2/m)^2}{m-1}+\frac{(s_2^2/n)^2}{n-1}}$$

$$\text{estimate}\pm(\text{c.v.})(\text{e.s.e.})$$

$$\frac{\text{estimate}-\text{param. value under }H_0}{\text{e.s.e. or (s.e. under }H_0)}$$

[3] **1.** The amount of aluminum concentration (measured in ppm) for a certain type of plastic was measured. The data from 10 samples are as follows:

119 30 76 182 222 79 87 143 118 101

Fill in each blank with the statistic asked for. It is only necessary to show work for part (ii).

(i) The sample mean, $\overline{x}$. 115.7

(ii) The sample median, $\widetilde{x}$. 109.5

(iii) The sample standard deviation, s. 55.5419

sorted: 30 76 79 87 101 118 119 143 182 222

$\tilde{x} = \frac{101+118}{2} = 109.5$

[3] **2.** Use the blank to indicate whether each statement is true or false. No reasons are necessary.

true (i) The sign of the covariance s_{xy} indicates the sign of the slope for the line of best fit between x and y. (i.e. positive or negative)

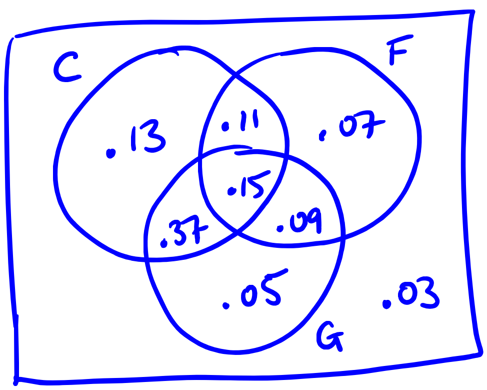
false (ii) A correlation coefficient of $r = 0$ indicates that there is no relationship between the variables in a sample of bivariate data.

true (iii) The correlation coefficient r indicates the strength of a linear relationship between x and y.

[3] **3.** A cafe keeps statistics on three characteristics: if a customer orders coffee, if a customer orders food, and if a customer takes their order to go. It is determined that:

- 76% order coffee C = coffee
- 42% order food F = food
- 66% take their order to go G = to go
- 26% order coffee and food
- 52% order coffee and take their order to go
- 24% order food and take their order to go
- 15% order coffee and food and take their order to go

Suppose a customer visiting the cafe is selected at random. What is the probability that they have exactly two of the characteristics? (For example, one such case is that the customer orders coffee and food but does not take their order to go.)



P(exactly two)

$= P(C \cap F \cap \overline{G}) + P(C \cap \overline{F} \cap G) + P(\overline{C} \cap F \cap G)$

$= 0.11 + 0.37 + 0.09$

$= 0.57$

Answer: 0.57

[3] **4.** A boiler has four identical relief valves. The probability that any particular valve will open on demand is 57%.
Assuming independent operation of the valves, what is the probability that at least one valve will open on demand?

P(at least one opens)

$= 1 - P(\text{none open})$

$= 1 - (1 - 0.57)^4$

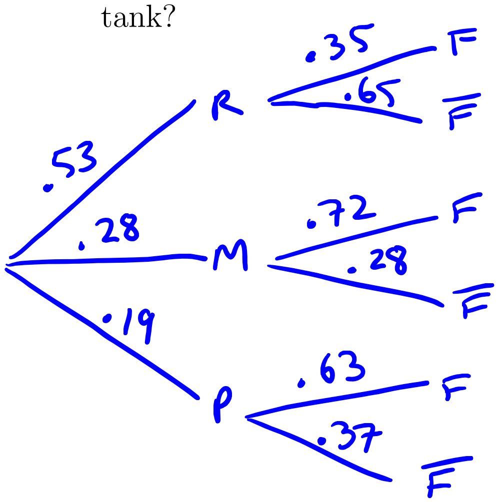
$= 1 - (0.43)^4$

$= 0.9658$

Answer: 0.9658

(Questions 5 and 6 refer to the following setup)
At a certain gas station customers can choose from three types of gas: regular, mid-grade, or premium. Suppose that 53% of customers choose regular, 28% choose mid-grade, and the rest choose premium gas.
Suppose further that of those who choose regular gas, 35% ask to completely fill their tank. Of those who choose mid-grade gas, 72% ask to completely fill their tank, and of those who choose premium gas, 63% ask to completely fill their tank.

[3] **5.** What is the probability that a randomly selected customer does not ask to completely fill their tank?



Answer

0.4932

$$P(\bar{F}) = (.53)(.65) + (.28)(.28) + (.19)(.37)$$
$$= 0.4932$$

[3] **6.** Suppose a customer who asks to completely fill their tank is selected at random. What is the probability that they purchase mid-grade gas?

Answer

0.3978

$$P(M \mid F) = \frac{P(M \cap F)}{P(F)}$$

$$= \frac{(.28)(.72)}{(.53)(.35) + (.28)(.72) + (.19)(.63)}$$

$$= \frac{0.2016}{0.5068} = 0.3978$$

(Questions 7 and 8 refer to the following setup)
While most shoppers now bring their own reusable bags when purchasing groceries, some shoppers have not adopted this habit (or occasionally shoppers will forget to bring their reusable bags) and need to purchase paper bags at the checkout. The number of paper bags each shopper purchases has the following cumulative distribution function:

	x	0	1	2	3	4	5	6	7
$P(X \le x) =$	$F(x)$	0.27	0.35	0.46	0.58	0.71	0.89	0.96	1
$P(X=x) =$	$f(x)$	.27	.08	.11	.12	.13	.18	.07	.04

[3] **7.** If it is known that a shopper purchases at least 3 paper bags, what is the probability that they purchase no more than 6 paper bags?

$$P(X \le 6 \mid X \ge 3) = \frac{P(3 \le X \le 6)}{P(X \ge 3)}$$

$$= \frac{P(X \le 6) - P(X \le 2)}{1 - P(X \le 2)} = \frac{0.96 - 0.46}{1 - 0.46} = \frac{0.50}{0.54} = 0.9259$$

or
$$\frac{P(3 \le X \le 6)}{P(X \ge 3)} = \frac{0.12 + 0.13 + 0.18 + 0.07}{0.12 + 0.13 + 0.18 + 0.07 + 0.04} = \frac{0.50}{0.54} = 0.9259$$

Answer
0.9259

[3] **8.** What is the expected number of paper bags a shopper purchases?

$$E(X) = \sum x \cdot f(x)$$
$$= 0(.27) + 1(.08) + 2(.11) + 3(.12)$$
$$+ 4(.13) + 5(.18) + 6(.07) + 7(.04)$$
$$= 2.78$$

Answer
2.78

(Question 9) A public transit system authority performs a study to investigate inefficiencies within the system. Here data is gathered on if buses arrive early, exactly on time, or late. The capacity of each bus is studied; they can be 1 - full, 2 - busy but not full, or 3 - rather empty. The proportion of each classification is summarized below:

E = early
T = on time
L = late

		Capacity			
		1	2	3	
Arrival	early	0.08	0.12	0.18	.38
	on time	0.04	0.15	0.07	.26
	late	0.21	0.09	0.06	.36
		.33	.36	.31	

(a) [**2 marks**] What is the probability that a randomly selected bus is full and will not be exactly on time?

$P(1 \cap \bar{T}) = 0.08 + 0.21 = 0.29$

(b) [**2 marks**] Are the events "bus is on time" and "bus is rather empty" independent? Support your answer with a calculation.

$P(T \cap 3) = 0.07$

$P(T) \cdot P(3) = (0.26)(0.31) = 0.0806$

$P(T \cap 3) \neq P(T) \cdot P(3)$ so the events are <u>not</u> independent.

The table from the previous page is copied below. (Recall that 1 - full, 2 - busy but not full, and 3 - rather empty.)

		Capacity			
		1	2	3	
Arrival	early	0.08	0.12	0.18	.38
	on time	0.04	0.15	0.07	.26
	late	0.21	0.09	0.06	.36
		.33	.36	.31	

(c) **[2 marks]** If a randomly selected bus is found to arrive early, what is the probability that it is full or busy but not full?

$$P(1 \cup 2 \mid E) = \frac{P((1 \cup 2) \cap E)}{P(E)}$$

$$= \frac{0.08 + 0.12}{0.08 + 0.12 + 0.18} = \frac{0.20}{0.38} = 0.5263$$

University of Victoria - Stat 260 - Spring 2023
Term Test 1 - Version B
Section A02 - Instructor: Dr. Michelle Edwards

Instructions:

- This paper has 6 pages of questions plus this cover page. Please count your pages and report any discrepancy immediately to the invigilator.
- Answer all questions on the test paper in the space provided. Questions 1 through 8 are short answer and worth 3 marks each. For each question show your work and write your final answer in the space as indicated. Question 9 is a full-answer question. To obtain full marks on this question, clear step-by-step reasoning must be displayed. Marks will be deducted for incoherent or incomplete solutions.
- Unless an exact answer with fewer decimals can be found, give all answers with at least three decimals of accuracy.
- A Sharp EL-510R series calculator may be used. No other calculator may be used.
- A formula sheet is provided. No other aids such as books, notes, scrap paper, electronic devices, etc. are permitted.
- You have 50 minutes to write the test.

Formula List

$$\frac{\sum_{i=1}^{n}(x_i-\overline{x})^2}{n-1}=\frac{\left(\sum_{i=1}^{n}x_i^2\right)-n\left(\overline{x}\right)^2}{n-1}$$

$$E(X^2)-\mu^2$$

$$\frac{\sum_{i=1}^{n}[(x_i-\bar{x})(y_i-\bar{y})]}{\sqrt{[\sum_{i=1}^{n}(x_i-\bar{x})^2][\sum_{i=1}^{n}(y_i-\bar{y})^2]}}$$

$$\frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}}$$

$$\binom{n}{x}p^x(1-p)^{n-x}, x=0,1,2,\ldots,n$$

$$\frac{\lambda^x}{x!}e^{-\lambda}, x=0,1,\ldots$$

$$1-e^{-\lambda x}, x>0$$

$$E(XY)-\mu_X\mu_Y$$

$$P\left(Z\leq\frac{x+0.5-np}{\sqrt{np(1-p)}}\right)$$

$$\frac{x^{\alpha-1}e^{-x/\beta}}{\beta^\alpha\Gamma(\alpha)}$$

$$\sum_{i=1}^{n}a_i^2\mathrm{Var}(X_i)+2\sum_{i<j}a_ia_j\mathrm{Cov}(X_i,X_j)$$

$$\frac{s}{\sqrt{n}}$$

$$\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{m}+\frac{\hat{p}_2(1-\hat{p}_2)}{n}}$$

$$\sqrt{\frac{s_1^2}{m}+\frac{s_2^2}{n}}$$

$$\sqrt{\frac{(m-1)s_1^2+(n-1)s_2^2}{m+n-2}\left(\frac{1}{m}+\frac{1}{n}\right)}$$

$$\text{integer part of }\frac{(s_1^2/m+s_2^2/n)^2}{\frac{(s_1^2/m)^2}{m-1}+\frac{(s_2^2/n)^2}{n-1}}$$

$$\text{estimate}\pm\text{(c.v.)(e.s.e.)}$$

$$\frac{\text{estimate}-\text{param. value under }H_0}{\text{e.s.e. or (s.e. under }H_0)}$$

[3] **1.** The amount of aluminum concentration (measured in ppm) for a certain type of plastic was measured. The data from 10 samples are as follows:

118 47 79 191 230 65 81 135 113 104

Fill in each blank with the statistic asked for. It is only necessary to show work for part (ii).

(i) The sample mean, $\overline{x}$. 116.3

(ii) The sample median, $\widetilde{x}$. 108.5

(iii) The sample standard deviation, s. 56.8566

sorted : 47 65 79 81 104 113 118 135 191 230

$$\tilde{x} = \frac{104+113}{2} = 108.5$$

[3] **2.** Use the blank to indicate whether each statement is true or false. No reasons are necessary.

true (i) A covariance of $s_{xy} = 0$ indicates there is no linear relationship between x and y.

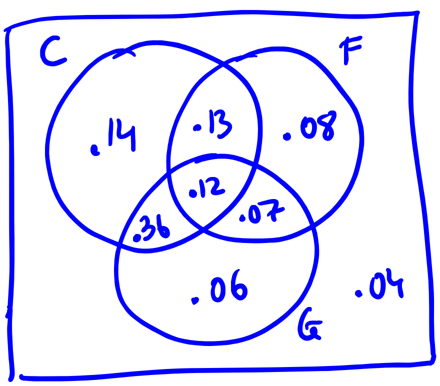
false (ii) The correlation coefficient r equals the slope of the straight line which best fits a sample of bivariate data.

false (iii) If the variables in a sample of bivariate data have a strong positive linear relation, then the correlation coefficient r is greater than 1.

[3] **3.** A cafe keeps statistics on three characteristics: if a customer orders coffee, if a customer orders food, and if a customer takes their order to go. It is determined that:

- 75% order coffee C = coffee
- 40% order food F = food
- 61% take their order to go G = to go
- 25% order coffee and food
- 48% order coffee and take their order to go
- 19% order food and take their order to go
- 12% order coffee and food and take their order to go

Suppose a customer visiting the cafe is selected at random. What is the probability that they have exactly two of the characteristics? (For example, one such case is that the customer orders coffee and food but does not take their order to go.)



P(exactly two)

$= P(C \cap F \cap \bar{G}) + P(C \cap \bar{F} \cap G)$

$+ P(\bar{C} \cap F \cap G)$

$= .13 + .36 + .07 = 0.56$

Answer

0.56

[3] **4.** A boiler has four identical relief valves. The probability that any particular valve will open on demand is 61%.
Assuming independent operation of the valves, what is the probability that at least one valve will open on demand?

P(at least one opens)

$= 1 - P(\text{none open})$

$= 1 - (1 - 0.61)^4$

$= 1 - (0.39)^4$

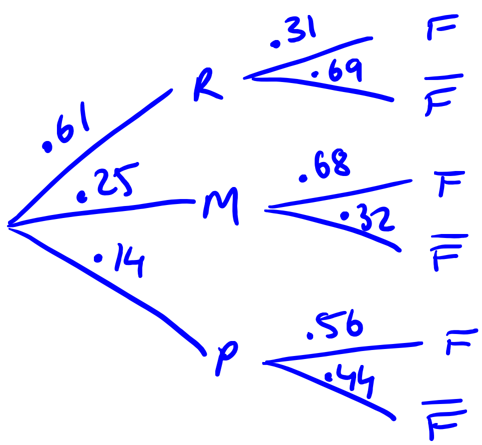
$= 0.9769$

Answer

0.9769

(Questions 5 and 6 refer to the following setup)
At a certain gas station customers can choose from three types of gas: regular, mid-grade, or premium. Suppose that 61% of customers choose regular, 25% choose mid-grade, and the rest choose premium gas.
Suppose further that of those who choose regular gas, 31% ask to completely fill their tank. Of those who choose mid-grade gas, 68% ask to completely fill their tank, and of those who choose premium gas, 56% ask to completely fill their tank.

[3] **5.** What is the probability that a randomly selected customer does not ask to completely fill their tank?



Answer
0.5625

$$P(\overline{F}) = (.61)(.69) + (.25)(.32) + (.14)(.44)$$
$$= 0.5625$$

[3] **6.** Suppose a customer who asks to completely fill their tank is selected at random. What is the probability that they purchase premium gas?

Answer
0.1792

$$P(P \mid F) = \frac{P(P \cap F)}{P(F)}$$
$$= \frac{(.14)(.56)}{(.61)(.31) + (.25)(.68) + (.14)(.56)}$$
$$= \frac{0.0784}{0.4375} = 0.1792$$

(Questions 7 and 8 refer to the following setup)
While most shoppers now bring their own reusable bags when purchasing groceries, some shoppers have not adopted this habit (or occasionally shoppers will forget to bring their reusable bags) and need to purchase paper bags at the checkout. The number of paper bags each shopper purchases has the following cumulative distribution function:

	x	0	1	2	3	4	5	6	7
$P(X \le x) =$	$F(x)$	0.31	0.44	0.62	0.72	0.80	0.89	0.95	1
$P(X=x) =$	$f(x)$	.31	.13	.18	.10	.08	.09	.06	.05

[3] **7.** If it is known that a shopper purchases at least 2 paper bags, what is the probability that they purchase no more than 5 paper bags?

$$P(X \le 5 \mid X \ge 2) = \frac{P(2 \le X \le 5)}{P(X \ge 2)}$$

$$= \frac{P(X \le 5) - P(X \le 1)}{1 - P(X \le 1)} = \frac{0.89 - 0.44}{1 - 0.44} = \frac{0.45}{0.56} = 0.8036$$

or $$\frac{P(2 \le X \le 5)}{P(X \ge 2)} = \frac{.18 + .10 + .08 + .09}{.18 + .10 + .08 + .09 + .06 + .05} = \frac{0.45}{0.56} = 0.8036$$

Answer
0.8036

[3] **8.** What is the expected number of paper bags a shopper purchases?

$$E(X) = \sum x \cdot f(x)$$

$$= 0(.31) + 1(.13) + 2(.18) + 3(.10)$$
$$+ 4(.08) + 5(.09) + 6(0.06) + 7(.05)$$

$$= 2.27$$

Answer
2.27

(Question 9) A public transit system authority performs a study to investigate inefficiencies within the system. Here data is gathered on if buses arrive early, exactly on time, or late. The capacity of each bus is studied; they can be 1 - full, 2 - busy but not full, or 3 - rather empty. The proportion of each classification is summarized below:

E = early

T = on time

L = late

		Capacity			
		1	2	3	
Arrival	early	0.06	0.08	0.16	.30
	on time	0.17	0.05	0.13	.35
	late	0.24	0.07	0.04	.35
		.47	.20	.33	

(a) [**2 marks**] What is the probability that a randomly selected bus is rather empty and will not be early?

$P(3 \cap \bar{E}) = 0.13 + 0.04 = 0.17$

(b) [**2 marks**] Are the events "bus is late" and "bus is busy but not full" independent? Support your answer with a calculation.

$P(L \cap 2) = 0.07$

$P(L) \cdot P(2) = (0.35)(0.20) = 0.07$

$P(L \cap 2) = P(L) \cdot P(2)$ so the events <u>are</u> independent.

The table from the previous page is copied below. (Recall that 1 - full, 2 - busy but not full, and 3 - rather empty.)

		Capacity			
		1	2	3	
Arrival	early	0.06	0.08	0.16	.30
	on time	0.17	0.05	0.13	.35
	late	0.24	0.07	0.04	.35
		.47	.20	.33	

(c) **[2 marks]** If a randomly selected bus is found to arrive exactly on time, what is the probability that it is busy but not full or rather empty?

$$P(2 \cup 3 \mid T) = \frac{P((2 \cup 3) \cap T)}{P(T)}$$

$$= \frac{0.05 + 0.13}{0.17 + 0.05 + 0.13} = \frac{0.18}{0.35} = 0.5143$$