國立政治大學 115 學年度第 1　學期　期中考　考試命題紙

考試科目：統計學 (一)　　開課班別：統計學整合開課　　命題教授：吳漢銘

考試日期：10/27(二)9:10–10:50	※准帶項目打「O」，否則打「×」					1. 需加發計算紙或答案紙請在試題內封袋備註。 2. 為環保節能減碳，試題一律採雙面印刷，有特殊印製需求，請註記：A 卷
本試題共 4 頁，印刷份數：145 份	計算機	課本	筆記	字典	手機平板筆電	
備註：注意事項要看!! (ch1∼5.4)	O	×	×	×	×	

注意事項:

(1) 在答題紙上填寫學號和姓名。請在答題紙右上角標明你使用的試卷編號：**A**（預設）或 **B**。

(2) 所有問題可用中文或英文作答（無需考慮語法和拼寫）。

(3) 請按題號順序作答。總分為 120 分。

(4) 建議使用深色原子筆（允許使用鉛筆）。禁用手機 3C 產品。可使用計算機 (無程式功能)。

(5) 計算過程（針對 **IV** 和 **V** 部分）必須寫出（計算至小數點後 4 位）。

(6) 答題紙與試題卷須一併交回。

(7) 作弊學生當次及日後考試試卷將不予批改，情節嚴重將報校處理。

(0) 宣誓詞 (0%): 複寫下列宣誓詞至答案卷的第一頁最上面。(**不寫扣 10 分**)

0. ” 本人姓名 恪遵各項考試規則，若如違反，願受校方最嚴厲處罰，謹誓。”

(I) 選擇題 (20%, 5% each); select one correct answer.

1. A researcher records the region of residence of 200 customers using the codes `1 = North, 2 = Central, 3 = South, 4 = East`. A student constructs a cumulative frequency distribution of the codes and reports that “the average region is 2.3.” Which statement is correct?
(A). The codes are numbers, so the data are quantitative and both summaries are meaningful.
(B). The data are categorical; neither summary is meaningful; use a bar chart or a pie chart instead.
(C). Only the cumulative frequency distribution is meaningful, because the codes can be ordered.
(D). A histogram is the most appropriate display because the codes are measured on an interval scale.

2. Of the last 100 customers who entered a computer shop, 40 purchased a computer. A student argues: “Each customer either buys or does not buy, so there are two experimental outcomes and, by the classical method, the probability that the next customer buys a computer is 1/2.” Which statement is correct?
(A). The student is correct; the classical method always assigns $1/n$ to each of the n experimental outcomes.
(B). The student is wrong: the classical method requires equally likely outcomes, which is not justified here; the relative frequency method gives the estimate 0.40.
(C). The student is wrong only because the sample space should be {buy, not buy, undecided}.
(D). Both 0.50 and 0.40 are acceptable, because any assignment with $0 \le P(E_i) \le 1$ and $\sum P(E_i) = 1$ is valid.

考試日期：10/27(二)9:10–10:50	※准帶項目打「O」，否則打「×」					1. 需加發計算紙或答案紙請在試題內封袋備註。 2. 為環保節能減碳，試題一律採雙面印刷，有特殊印製需求，請註記：A 卷
本試題共 4 頁，印刷份數: 145 份	計算機	課本	筆記	字典	手機平板筆電	
備註：注意事項要看!! (ch1∼5.4)	O	×	×	×	×	

3. Let A and B be two events with $P(A) = 0.6$ and $P(B) = 0.5$. Which of the following statements **must** be true?
   (A). A and B cannot be mutually exclusive. (B). A and B must be independent.
   (C). $P(A \cup B) = 0.8$. (D). $P(A \cap B) = 0.3$.
4. For which of the following is $f(x)$ a valid probability distribution for the discrete random variable x?
   (A). $f(x) = (x-2)/2$ for $x = 1, 2, 3$. (B). $f(x) = (3-x)/6$ for $x = 0, 1, 2$.
   (C). $f(x) = (x+1)/8$ for $x = 1, 2, 3$. (D). $f(x) = x/4$ for $x = 1, 2, 3$.

(II) 填空題 (Correct spelling should be used.) (10%, 5% each; 2.5% per blank)

5. The sample correlation coefficient r_{xy} is a measure of the _____ association between x and y; a value of r_{xy} close to $+1$ or -1 indicates a strong relationship, but it is not necessarily evidence of _____.

6. Formally, a random variable X is a _____ that assigns a real number to each experimental outcome. When a large amount of historical data is treated as if it were the population and the relative frequency method is used to assign probabilities to the values of X, the resulting probability distribution is called an _____ discrete distribution.

(III) 簡答題 (20%, 10% each) (write down the statement (or definition), formula if any, interpretation)

7. The textbook classifies the scale of measurement of a variable as nominal, ordinal, interval, or ratio. Explain the difference between the **interval** scale and the **ratio** scale. Your answer must state the property that the ratio scale has but the interval scale does not, and the role of the value zero.

7. Let x and y denote the percent returns of a stock fund and a bond fund, respectively, with expected values $E(x)$, $E(y)$, variances $\mathrm{Var}(x)$, $\mathrm{Var}(y)$, and covariance σ_{xy}. An investor puts a proportion a of her money in the stock fund and the remaining proportion $b = 1 - a$ in the bond fund, so that the percent return of the portfolio is $r = ax + by$.

   (a). (5%) Write down the formulas for $E(r)$ and $\mathrm{Var}(r)$ in terms of $E(x)$, $E(y)$, $\mathrm{Var}(x)$, $\mathrm{Var}(y)$, σ_{xy}, a, and b.

   (b). (5%) Using the formula for $\mathrm{Var}(r)$, explain why the standard deviation (risk) of the portfolio can be **smaller** than the standard deviations of **both** funds. Which sign of σ_{xy} makes this risk reduction strongest? Why?

(IV) 計算題 (50%, 25% each) (計算至小數點後 4 位)

8. **On-Time Performance of Airlines.** Three airlines share one terminal of an airport. Among all flights arriving at this terminal, 25% are flights of airline A, 35% are flights of airline B, and 40% are flights of airline C. The percentages of on-time arrivals are 80% for airline A, 70% for airline B, and 85% for airline C. A flight arriving at this terminal is selected at random. Let O denote the event that the flight arrives on time, and L the event that it arrives late.

   (a). (5%) Develop the joint probability table with three rows (the airlines) and two columns (O and L), including the marginal probabilities.

   (b). (5%) What is the probability that the selected flight arrives on time?

   (c). (10%) Before any information about the arrival is known, which airline is the most likely for the selected flight? Now an announcement says that the flight will arrive late. Given this information, which airline is the most likely for the flight, and what is the (posterior) probability that the flight belongs to this airline?

   (d). (5%) Are the events "the flight belongs to airline C" and O independent? Justify your answer using probabilities.

9. **Printer Manufacturing Costs.** PortaCom has developed a design for a high-quality portable printer. The two key components of the manufacturing cost per printer are the parts cost and the direct labor cost. Let x = parts cost (in dollars) per printer and y = direct labor cost (in dollars) per printer. Based on extensive product tests, the engineers developed the following bivariate probability distribution $f(x, y)$ for the two costs.

	Direct Labor Cost y		
Parts Cost x	40	45	50
80	0.20	0.25	0.05
90	0.10	0.15	0.25

(a). (10%) Show the marginal probability distributions of x and of y, and compute $E(x)$, $\text{Var}(x)$, $E(y)$, and $\text{Var}(y)$.

(b). (5%) Let $t = x + y$ be the total manufacturing cost per printer. Show the probability distribution of t, and compute $E(t)$ and $\text{Var}(t)$. (Hint: t takes five possible values.)

(c). (5%) Compute the covariance σ_{xy} (use $\text{Var}(x + y)$, $\text{Var}(x)$, and $\text{Var}(y)$) and the correlation coefficient ρ_{xy} between x and y. What do they tell you about the relationship between the parts cost and the direct labor cost?

(d). (5%) Are x and y independent? Justify your answer using probabilities.

(V) 加分題 (20%) (Proofs are required; a numerical example alone is not a proof.)

10. Let A and B be two events in a sample space $\mathcal{S}$ with $P(A) > 0$ and $P(B) > 0$, and that A^c denotes the complement of A.

(a). (10%) Suppose that A and B are independent. Prove that A and B^c are also independent.

(b). (10%) Suppose instead that A and B are mutually exclusive. Prove that A and B must be **dependent**. Then compute $P(A \mid B)$ and explain, in one sentence, what this value tells you about A once B is known to have occurred.

<The blank pages at the back can be used as scratch paper. (後面空白頁可當計算紙)>

注意： 1、考試求公平及公正，請同學務必自律，維護學校與學生之榮譽。

2、考試時不得有交談、窺視、夾帶、抄襲、傳遞、代考或其它作弊等舞弊行為，考畢務必交卷，不得攜卷出場，違者依考場規則議處。