Week 14, Problem Set 2 - worked solutions

Wackerly 5.9 - The Multinomial Probability Distribution

> This is one randomly generated version of the set. Every student receives different numbers, so the values here will not match your own copy — follow the method, not the arithmetic.

Problem 1

seed 20260925

A trader at a Baku brokerage who hedges a jeweller's gold purchases classifies each daily London close of the gold price as **up** (more than 0.5% above the previous close), **flat** (within 0.5%) or **down** (more than 0.5% below). She models the days as independent, with

$$P(\text{up}) = 0.44, \qquad P(\text{flat}) = 0.25, \qquad P(\text{down}) = 0.31$$

on every day. Over the next $n = 7$ trading days let Y_1, Y_2, Y_3 be the numbers of up, flat and down days.

Give each answer as a decimal to at least four significant figures.

(a) $P(Y_1 = 4,\ Y_2 = 1,\ Y_3 = 2) =$ ______________

(b) The expected number of up days, $E(Y_1) =$ ______________

(c) The variance of the number of down days, $V(Y_3) =$ ______________

The $n = 7$ days are identical, independent trials, each falling into one of $k = 3$ cells with fixed probabilities, so Definition 5.11 is satisfied and (Y_1, Y_2, Y_3) is multinomial with parameters $n = 7$ and $p_1 = 0.44$, $p_2 = 0.25$, $p_3 = 0.31$ (Definition 5.12).

(a) The counts add to $4 + 1 + 2 = 7$, as they must.

$$p(4, 1, 2) = \frac{7!}{4!\,1!\,2!}(0.44)^4(0.25)^1(0.31)^2 = 105\,(0.44)^4(0.25)^1(0.31)^2 = 0.094550.$$

The coefficient 105 counts the orderings of the 7 days that give this pattern; each ordering has the same probability, the product of powers.

(b) By Theorem 5.13, $E(Y_1) = np_1 = 7(0.44) = 3.08$.

(c) By Theorem 5.13, $V(Y_3) = np_3q_3 = 7(0.31)(0.69) = 1.4973$. Lumping up and flat days into one cell, Y_3 is binomial with $n = 7$ and $p = 0.31$, which is where this formula comes from.

Problem 2

seed 20260925

A mobile operator's shop in Ganja signs up $n = 11$ new prepaid subscribers on a typical day. Each chooses one of three tariffs, independently of the others, with probabilities

$$\textbf{Mini: } p_1 = 0.39, \qquad \textbf{Standard: } p_2 = 0.44, \qquad \textbf{Unlimited: } p_3 = 0.17.$$

Let Y_1, Y_2, Y_3 be the numbers of the day's subscribers who choose Mini, Standard and Unlimited. Give each answer to at least four significant figures.

(a) $E(Y_1) =$ ______________

(b) $V(Y_2) =$ ______________

(c) The standard deviation of Y_3 is ______________

(d) $\text{Cov}(Y_1, Y_3) =$ ______________

The day's sign-ups are $n = 11$ independent trials, each landing in one of three cells with fixed probabilities, so (Y_1, Y_2, Y_3) is multinomial (Definition 5.12) and Theorem 5.13 applies.

(a) $E(Y_1) = np_1 = 11(0.39) = 4.29$.

(b) $V(Y_2) = np_2q_2 = 11(0.44)(0.56) = 2.7104$.

(c) $V(Y_3) = np_3q_3 = 11(0.17)(0.83) = 1.5521$, so $\sigma_{Y_3} = \sqrt{1.5521} = 1.2458$.

(d) By part 2 of Theorem 5.13, $\mathrm{Cov}(Y_1, Y_3) = -np_1p_3 = -11(0.39)(0.17) = -0.7293$.

The sign is forced by the fixed total: the 11 subscribers have to go somewhere, so a day with more Mini sign-ups leaves fewer customers to choose Unlimited.

Problem 3

seed 20260925

A bank's corporate book contains $n = 9$ firms rated BBB at the start of the year. From the rating agency's transition matrix, each firm independently ends the year upgraded (to A or better) with probability 0.14, still BBB with probability 0.77, or downgraded (to BB or worse) with probability 0.09. Let Y_1, Y_2, Y_3 be the numbers of firms upgraded, unchanged and downgraded.

Give each number to at least four significant figures.

(a) $\mathrm{Cov}(Y_1, Y_3) =$ ____________

(b) $\mathrm{Cov}(Y_1, Y_2) =$ ____________

(c) The correlation coefficient ρ between Y_1 and Y_3 is ____________

(d) Which pair of counts has the most negative covariance?

- ○ Upgrades and downgrades
- ○ Upgrades and firms staying BBB
- ○ Downgrades and firms staying BBB

(Y_1, Y_2, Y_3) is multinomial with $n = 9$, $p_1 = 0.14$, $p_2 = 0.77$, $p_3 = 0.09$, and part 2 of Theorem 5.13 gives $\mathrm{Cov}(Y_s, Y_t) = -np_sp_t$.

(a) $\mathrm{Cov}(Y_1, Y_3) = -9(0.14)(0.09) = -0.1134$.

(b) $\mathrm{Cov}(Y_1, Y_2) = -9(0.14)(0.77) = -0.9702$.

(c) $V(Y_1) = 9(0.14)(0.86) = 1.0836$ and $V(Y_3) = 9(0.09)(0.91) = 0.7371$, so

$$\rho = \frac{\mathrm{Cov}(Y_1, Y_3)}{\sqrt{V(Y_1)V(Y_3)}} = \frac{-0.1134}{\sqrt{(1.0836)(0.7371)}} = -0.1269.$$

The n cancels: $\rho = -\sqrt{p_1p_3/(q_1q_3)}$, whatever the size of the pool.

(d) All three covariances are $-n$ times a product of two probabilities, so the most negative one belongs to the pair with the largest product. The third covariance is $\mathrm{Cov}(Y_2, Y_3) = -0.6237$. Since "stay BBB" is by far the largest cell, the winner pairs it with the larger of the two moving cells, here the upgrade cell. The answer is: Upgrades and firms staying BBB.

Problem 4

seed 20260925

An insurer in Baku covers a delivery company's fleet of $n = 9$ vans under separate motor policies. Over a year each policy, independently, ends up in one of three classes: **claim-free** (probability 0.74), **small claims only** (probability 0.16), or **at least one large claim** (probability 0.10). Let Y_1, Y_2, Y_3 be the numbers of policies in the three classes.

Give each answer as a decimal to at least four significant figures.

(a) The probability that no policy has a large claim, $P(Y_3 = 0) =$ ____________

(b) $P(Y_2 = 3) =$ ____________

(c) $P(Y_3 \geq 2) =$ ____________

(d) The expected number of claim-free policies, $E(Y_1) =$ ____________

(Y_1, Y_2, Y_3) is multinomial with $n = 9$ and cell probabilities $0.74, 0.16, 0.10$. As in the proof of Theorem 5.13, merge every cell except cell i into one: each policy is then "in cell i" or "not", with probabilities p_i and $1 - p_i$, so Y_i is binomial with parameters n and p_i. No sum over the other counts is needed.

(a) $Y_3 \sim \text{binomial}(9, 0.10)$, so $P(Y_3 = 0) = (0.90)^9 = 0.387420$.

(b) $Y_2 \sim \text{binomial}(9, 0.16)$, so

$$P(Y_2 = 3) = \binom{9}{3}(0.16)^3(0.84)^6 = 84(0.16)^3(0.84)^6 = 0.120869.$$

(c) Use the complement: $P(Y_3 = 1) = 9(0.10)(0.90)^8 = 0.387420$, so $P(Y_3 \geq 2) = 1 - 0.387420 - 0.387420 = 0.225159$.

(d) $E(Y_1) = np_1 = 9(0.74) = 6.66$.

Problem 5

seed 20260925

A microfinance organisation in Lankaran disburses $n = 8$ new loans this month. Three months later each loan, independently, is in one of four states:

- current, with probability $p_1 = 0.72$;
- 1 to 30 days late, with probability $p_2 = 0.17$;
- 31 to 90 days late, with probability $p_3 = 0.06$;
- in default, with probability $p_4 = 0.05$.

Let Y_1, Y_2, Y_3, Y_4 be the numbers of loans in the four states. Give each answer to at least four significant figures.

(a) The probability that exactly 2 loans are late (in either lateness bucket) and exactly 1 are in default, i.e. $P(Y_2 + Y_3 = 2,\ Y_4 = 1) =$

____

(b) $P(Y_2 + Y_3 = 2) =$ ______________

(c) $E(Y_2 + Y_3) =$ ______________

Merging cells does not break Definition 5.11: each loan still falls into exactly one of the new, coarser cells, with a fixed probability, independently of the others. So the counts of the merged cells are again multinomial, with the merged probabilities.

(a) Put the two lateness buckets together. With $L = Y_2 + Y_3$, the vector (Y_1, L, Y_4) is multinomial with $n = 8$ and probabilities 0.72, $0.17 + 0.06 = 0.23$, 0.05. The event asks for $Y_1 = 8 - 2 - 1 = 5$, $L = 2$, $Y_4 = 1$:

$$\frac{8!}{5!\,2!\,1!}(0.72)^5(0.23)^2(0.05)^1 = 168 \times (0.72)^5(0.23)^2(0.05)^1 = 0.085980.$$

(b) Now merge everything except "late": L is binomial with $n = 8$, $p = 0.23$:

$$P(L = 2) = \binom{8}{2}(0.23)^2(0.77)^6 = 28(0.23)^2(0.77)^6 = 0.308715.$$

(c) $E(L) = n(p_2 + p_3) = 8(0.23) = 1.84$, the same as $E(Y_2) + E(Y_3)$ by Theorem 5.13 and linearity.

Problem 6

seed 20260925

A coffee chain in Baku records how customers pay. For a randomly chosen customer the probabilities are

Method	Cash	Bank card	Mobile wallet	Loyalty points
Probability	0.36	0.39	0.14	0.11

Customers choose independently. The next $n = 6$ customers are observed and Y_1, Y_2, Y_3, Y_4 count the payments by cash, card, wallet and points.

Give each answer as a decimal to at least four significant figures.

(a) $P(Y_1 = 1,\ Y_2 = 2,\ Y_3 = 2,\ Y_4 = 1) =$ ______________

(b) The probability that none of the 6 customers pays with the mobile wallet is ______________

The 6 customers are identical independent trials with four outcome cells, so (Y_1, Y_2, Y_3, Y_4) is multinomial (Definition 5.12) with $n = 6$ and $p_1 = 0.36$, $p_2 = 0.39$, $p_3 = 0.14$, $p_4 = 0.11$.

(a) The counts add to 6, so the formula applies directly (compare Example 5.30, where a cell with count zero contributes $0! = 1$ and $p^0 = 1$):

$$p(1,2,2,1) = \frac{6!}{1!\,2!\,2!\,1!}(0.36)^1(0.39)^2(0.14)^2(0.11)^1 = 180 \times (0.36)^1(0.39)^2(0.14)^2(0.11)^1 = 0.021250.$$

(b) The marginal of Y_3 is binomial(6, 0.14), so $P(Y_3 = 0) = (0.86)^6 = 0.404567$.

Problem 7

seed 20260925

A commercial bank's branch network opens $n = 28$ new deposit accounts in a week. Each new customer, independently, chooses a demand deposit (probability 0.39), a 6-month term deposit (probability 0.39) or a 12-month term deposit (probability 0.22). Let Y_1, Y_2, Y_3 be the numbers of accounts of each kind.

The treasury watches $D = Y_2 - Y_3$, the excess of 6-month over 12-month accounts, because it shifts the maturity profile of the bank's funding.

Give each answer to at least four significant figures.

(a) $\mathrm{Cov}(Y_2, Y_3) =$ ______________

(b) $E(D) =$ ______________

(c) $V(D) =$ ______________

(d) The standard deviation of D is ______________

(Y_1, Y_2, Y_3) is multinomial with $n = 28$ and $p = (0.39, 0.39, 0.22)$.

(a) Theorem 5.13: $\mathrm{Cov}(Y_2, Y_3) = -np_2p_3 = -28(0.39)(0.22) = -2.4024$.

(b) $D = a_1Y_2 + a_2Y_3$ with $a_1 = 1$, $a_2 = -1$. By Theorem 5.12, $E(D) = np_2 - np_3 = 28(0.17) = 4.76$.

(c) Theorem 5.13 gives $V(Y_2) = 28(0.39)(0.61) = 6.6612$ and $V(Y_3) = 28(0.22)(0.78) = 4.8048$. Theorem 5.12 then gives

$$V(D) = (1)^2V(Y_2) + (-1)^2V(Y_3) + 2(1)(-1)\mathrm{Cov}(Y_2, Y_3) = 6.6612 + 4.8048 + 4.8048 = 16.2708.$$

(d) $\sigma_D = \sqrt{16.2708} = 4.0337$.

The covariance term matters: ignoring it would give 11.4660, too small. Because the two counts move in opposite directions, their difference swings more than either count alone would suggest.

Problem 8

seed 20260925

A bank's collections unit receives $n = 26$ newly overdue credit-card accounts each week. Independently, each account needs a reminder phone call (probability 0.54), needs a field visit by an agent (probability 0.13), or is paid by the customer without any contact (probability 0.33). Let Y_1 and Y_2 be the numbers of accounts needing a call and a visit.

A call costs 6 AZN and a visit costs 28 AZN, so the week's collection cost is $C = 6Y_1 + 28Y_2$ AZN.

Give each answer to at least four significant figures.

(a) $E(C) =$ ______________ AZN

(b) $V(C) =$ ______________ AZN^2

(c) The standard deviation of C is ______________ AZN

With the third cell ("no contact", Y_3) included, (Y_1, Y_2, Y_3) is multinomial with $n = 26$ and $p = (0.54, 0.13, 0.33)$. Theorem 5.13 gives

$$E(Y_1) = 14.04, \quad E(Y_2) = 3.38, \quad V(Y_1) = 26(0.54)(0.46) = 6.4584, \quad V(Y_2) = 26(0.13)(0.87) = 2.9406,$$
$$\mathrm{Cov}(Y_1, Y_2) = -26(0.54)(0.13) = -1.8252.$$

(a) Theorem 5.12(a): $E(C) = 6(14.04) + 28(3.38) = 178.88$ AZN.

(b) Theorem 5.12(b) with $a_1 = 6$, $a_2 = 28$:

$$V(C) = 36\,V(Y_1) + 784\,V(Y_2) + 336\,\mathrm{Cov}(Y_1, Y_2) = 36(6.4584) + 784(2.9406) + 336(-1.8252) = 1924.6656.$$

(c) $\sigma_C = \sqrt{1924.6656} = 43.8710$ AZN.

The negative covariance lowers the variance: a week with many accounts needing calls is, for a fixed n, a week with fewer accounts left over to need visits.

Problem 9

seed 20260925

A credit fund holds bonds of $n = 11$ issuers rated BB. Over a year each issuer, independently, is upgraded to investment grade (probability 0.17), stays BB (probability 0.68) or is downgraded to B or worse (probability 0.15). Let Y_1, Y_2, Y_3 count the upgraded, unchanged and downgraded issuers.

At year end the risk report arrives first with only the downgrades: $Y_3 = 4$. Given this information, answer the following to at least four significant figures.

(a) For one of the issuers **not** downgraded, the conditional probability that it was upgraded is ______________

(b) $P(Y_1 = 1 \mid Y_3 = 4) =$ ______________

(c) $E(Y_1 \mid Y_3 = 4) =$ ______________

(d) $V(Y_1 \mid Y_3 = 4) =$ ______________

(a) For a single issuer, Definition 2.9 gives $P(\text{up} \mid \text{not down}) = p_1/(p_1 + p_2) = 0.17/0.85 = 0.200000$.

(b) By Definition 5.12 and the binomial marginal of Y_3,

$$P(Y_1 = y_1 \mid Y_3 = 4) = \frac{p(y_1, 7 - y_1, 4)}{P(Y_3 = 4)} = \frac{\frac{11!}{y_1!(7-y_1)!\,4!} p_1^{y_1} p_2^{7-y_1} p_3^4}{\frac{11!}{4!\,7!}(p_1 + p_2)^7 p_3^4} = \binom{7}{y_1} \left(\frac{p_1}{p_1 + p_2}\right)^{y_1} \left(\frac{p_2}{p_1 + p_2}\right)^{7-y_1}.$$

So given $Y_3 = 4$, Y_1 is binomial with 7 trials and success probability 0.200000. The 7 issuers not downgraded are independent, and each was upgraded with the probability in (a). Therefore

$$P(Y_1 = 1 \mid Y_3 = 4) = \binom{7}{1}(0.200000)^1(0.800000)^6 = 0.367002.$$

(c) The binomial mean is $E(Y_1 \mid Y_3 = 4) = 7(0.200000) = 1.4000$.

(d) The binomial variance is $V(Y_1 \mid Y_3 = 4) = 7(0.200000)(0.800000) = 1.1200$.

Compare $E(Y_1) = 1.8700$ and $P(Y_1 = 1) = 0.290150$ before the report arrived. Learning the number of downgrades changes what we expect about upgrades. That is the dependence that Theorem 5.13's non-zero covariance describes.

Problem 10

seed 20260925

A retail client of a Baku broker places $n = 11$ limit orders on a foreign stock through the broker's app during a week. Each order, independently, ends up as an executed buy (probability 0.54), an executed sell (probability 0.30) or cancelled before execution (probability 0.16). Let Y_1, Y_2, Y_3 be the numbers of executed buys, executed sells and cancellations.

Give each answer as a decimal to at least four significant figures.

(a) The probability of at least 3 cancellations, $P(Y_3 \geq 3) =$ ______________

(b) $P(Y_1 = 5,\ Y_3 = 0) =$ ____________

(c) $P(Y_1 = 5 \mid Y_3 = 0) =$ ____________

(Y_1, Y_2, Y_3) is multinomial with $n = 11$ and $p = (0.54, 0.30, 0.16)$.

(a) The marginal of Y_3 is binomial$(11, 0.16)$ (proof of Theorem 5.13). Then $P(Y_3 = 0) = (0.84)^{11} = 0.146917$, $P(Y_3 = 1) = 11(0.16)(0.84)^{11-1} = 0.307826$ and $P(Y_3 = 2) = \binom{11}{2}(0.16)^2(0.84)^{11-2} = 55(0.16)^2(0.84)^{11-2} = 0.293168$. Using the complement over $y_3 = 0, \ldots, 2$, $P(Y_3 \geq 3) = 1 - (0.146917 + 0.307826 + 0.293168) = 0.252089$.

(b) $Y_3 = 0$ and $Y_1 = 5$ force $Y_2 = 6$. Definition 5.12 gives

$$p(5, 6, 0) = \frac{11!}{5!\,6!\,0!}(0.54)^5(0.30)^6(0.16)^0 = 462(0.54)^5(0.30)^6 = 0.015465.$$

(c) Definition 2.9: $0.015465/0.146917 = 0.105261$. As a check, given no cancellations all 11 orders are buys or sells. Each is a buy with probability $0.54/(0.54 + 0.30) = 0.642857$, so the answer is also $\binom{11}{5}(0.642857)^5(0.357143)^6$, the same number.

Problem 11

seed 20260925

A currency-exchange kiosk at Baku's Heydar Aliyev airport serves customers who each, independently, buy US dollars (probability 0.44), sell US dollars (probability 0.28) or make a transaction in another currency (probability 0.28). Among the next $n = 4$ customers let Y_1, Y_2, Y_3 be the numbers in each group. The cashier's dollar float falls if more customers buy dollars than sell them.

Give each answer as a decimal to at least four significant figures.

(a) $P(Y_1 > Y_2) =$ ____________

(b) $P(Y_1 = Y_2) =$ ____________

(Y_1, Y_2, Y_3) is multinomial with $n = 4$ and $p = (0.44, 0.28, 0.28)$. You can add Definition 5.12 over every cell with $y_1 > y_2$. It is shorter to split the problem into two binomial steps.

Step 1 (grouping). Let $S = Y_1 + Y_2$ be the number of dollar customers. Merging the first two cells, S is binomial with $n = 4$ and $p = 0.44 + 0.28 = 0.72$.

Step 2 (conditioning). Given $S = s$, each of those s customers is, independently, a buyer with probability $0.44/0.72 = 0.611111$. So $Y_1 \mid S = s$ is binomial$(s, 0.611111)$. The same cancellation as in the conditional multinomial calculation gives this.

Then $P(Y_1 > Y_2) = \sum_s P(S = s)\, P(Y_1 > s/2 \mid S = s)$, and similarly for $Y_1 = Y_2$ (possible only for even s):

s	$P(S = s)$	$P(Y_1 > s/2 \mid S = s)$	$P(Y_1 = s/2 \mid S = s)$
0	0.006147	0.000000	1.000000
1	0.063222	0.611111	0.000000
2	0.243855	0.373457	0.475309
3	0.418038	0.663923	0.000000
4	0.268739	0.494484	0.338877

(a) Multiplying columns 2 and 3 and adding gives $P(Y_1 > Y_2) = 0.540137$.

(b) Multiplying columns 2 and 4 and adding gives $P(Y_1 = Y_2) = 0.213123$. This includes $s = 0$, the case with no dollar customers at all.

Hence $P(Y_1 < Y_2) = 1 - 0.540137 - 0.213123 = 0.246740$.

Problem 12

seed 20260925

A mortgage-backed bond issued by a Baku bank is backed by a pool of $n = 56$ residential mortgages. Over the coming year each mortgage, independently, is prepaid in full (probability 0.14), defaults (probability 0.04) or stays current (probability 0.82). Let Y_1, Y_2, Y_3 be the counts, and let $N = Y_1 + Y_2$ be the number of mortgages that leave the pool, which is what shortens the bond's life.

Give each answer to at least four significant figures.

(a) $E(N) =$ ______________

(b) $\text{Cov}(Y_1, Y_2) =$ ______________

(c) $V(N) =$ ______________

(d) The correlation coefficient between Y_1 and Y_2 is ______________

(Y_1, Y_2, Y_3) is multinomial with $n = 56$ and $p = (0.14, 0.04, 0.82)$.

(a) $E(N) = np_1 + np_2 = 56(0.18) = 10.08$.

(b) Theorem 5.13: $\text{Cov}(Y_1, Y_2) = -np_1p_2 = -56(0.14)(0.04) = -0.3136$.

(c) By Theorem 5.13, $V(Y_1) = 56(0.14)(0.86) = 6.7424$ and $V(Y_2) = 56(0.04)(0.96) = 2.1504$. Theorem 5.12 with $a_1 = a_2 = 1$ gives

$$V(N) = V(Y_1) + V(Y_2) + 2\text{Cov}(Y_1, Y_2) = 6.7424 + 2.1504 + (-0.6272) = 8.2656.$$

Check by grouping. Merging the prepay and default cells, N is binomial with $n = 56$ and $p = 0.18$, so $V(N) = 56(0.18)(0.82) = 8.2656$. The two methods agree, as they must: $np_1q_1 + np_2q_2 - 2np_1p_2 = n(p_1 + p_2)(1 - p_1 - p_2)$.

(d) $\rho = -0.3136/\sqrt{(6.7424)(2.1504)} = -0.0824$. The correlation is negative but weak, because both cells are small compared with "stays current".

Problem 13

seed 20260925

A consultancy surveys $n = 240$ randomly chosen Baku households and asks which bank is their main bank. Suppose the true shares are $p_1 = 0.41$ for Bank A, $p_2 = 0.20$ for Bank B and $p_3 = 0.39$ for all other banks together. Treat the answers as independent. Let Y_1, Y_2, Y_3 be the numbers of households naming Bank A, Bank B and another bank.

Bank A's marketing team will quote the estimated lead $\hat{L} = \dfrac{Y_1}{n} - \dfrac{Y_2}{n}$.

Give each answer to at least four significant figures.

(a) $E(\hat{L}) =$ ______________

(b) $V(\hat{L}) =$ ______________

(c) The standard deviation of $\hat{L}$ is ______________

(Y_1, Y_2, Y_3) is multinomial with $n = 240$ and $p = (0.41, 0.20, 0.39)$. Write $\hat{L} = a_1Y_1 + a_2Y_2$ with $a_1 = 1/n$ and $a_2 = -1/n$.

(a) Theorems 5.12 and 5.13: $E(\hat{L}) = (np_1 - np_2)/n = p_1 - p_2 = 0.21$. The estimated lead is unbiased for the true lead.

(b) Theorem 5.12(b):

$$V(\hat{L}) = \frac{1}{n^2}V(Y_1) + \frac{1}{n^2}V(Y_2) - \frac{2}{n^2}\text{Cov}(Y_1, Y_2) = \frac{np_1q_1 + np_2q_2 + 2np_1p_2}{n^2} = \frac{p_1q_1 + p_2q_2 + 2p_1p_2}{n}.$$

With the numbers, $(0.41)(0.59) + (0.20)(0.80) + 2(0.41)(0.20) = 0.5659$, so $V(\hat{L}) = 0.5659/240 = 0.00235792$.

(c) $\sqrt{0.00235792} = 0.048558$.

If the negative covariance is ignored, the standard deviation comes out as 0.040922, which understates the error. A household that does not name Bank A is more likely to name Bank B, so the two estimated shares move in opposite directions and their difference is noisier.

Problem 14

seed 20260925

A development bank funds a cohort of $n = 11$ small-business loans with the same term. At maturity each loan, independently, has been repaid early (probability 0.29), repaid on schedule (0.49), restructured (0.11) or has defaulted (0.11). Let Y_1, Y_2, Y_3, Y_4 count the four outcomes.

The workout department reports first: exactly $Y_4 = 2$ loans defaulted. Given this, answer the following to at least four significant figures.

(a) $P(Y_1 = 3,\ Y_2 = 4,\ Y_3 = 2 \mid Y_4 = 2) =$ ______________

(b) $E(Y_3 \mid Y_4 = 2) =$ ______________

Write the conditional probability as a ratio (Definition 2.9). The numerator is Definition 5.12. The denominator is the binomial marginal of Y_4 :

$$\frac{\frac{11!}{y_1!\,y_2!\,y_3!\,2!}p_1^{y_1}p_2^{y_2}p_3^{y_3}p_4^2}{\frac{11!}{2!\,9!}(1-p_4)^9p_4^2} = \frac{9!}{y_1!\,y_2!\,y_3!}\left(\frac{p_1}{1-p_4}\right)^{y_1}\left(\frac{p_2}{1-p_4}\right)^{y_2}\left(\frac{p_3}{1-p_4}\right)^{y_3},$$

for $y_1 + y_2 + y_3 = 9$.

So, given $Y_4 = 2$, (Y_1, Y_2, Y_3) is multinomial with 9 trials and probabilities $0.29/0.89 = 0.325843$, $0.49/0.89 = 0.550562$ and $0.11/0.89 = 0.123596$. The non-defaulted loans are still independent trials, now with the default cell removed.

(a)

$$\frac{9!}{3!\,4!\,2!}(0.325843)^3(0.550562)^4(0.123596)^2 = 1260 \times (0.325843)^3(0.550562)^4(0.123596)^2 = 0.061182.$$

(b) Theorem 5.13 for the conditional multinomial: $E(Y_3 \mid Y_4 = 2) = 9(0.123596) = 1.1124$. Compare the unconditional $E(Y_3) = 11(0.11) = 1.2100$.

Problem 15

seed 20260925

A telecom operator's call centre handles n contract renewals a day. Each customer, independently, picks a premium bundle, a standard bundle or a basic bundle, with fixed probabilities p_1, p_2, p_3. The numbers Y_1, Y_2, Y_3 choosing each bundle are therefore multinomial. The analytics team does not report n or the p_i. It reports only

$$E(Y_1) = 10.24, \qquad V(Y_1) = 6.9632, \qquad E(Y_2) = 12.16.$$

Give each answer to at least four significant figures.

(a) $p_1 =$ ______________

(b) $n =$ ______________

(c) $p_2 =$ ______________

(d) $\mathrm{Cov}(Y_1, Y_2) =$ ______________

By Theorem 5.13, $E(Y_1) = np_1$ and $V(Y_1) = np_1q_1$.

(a) Divide the two: $V(Y_1)/E(Y_1) = q_1 = 6.9632/10.24 = 0.68$, so $p_1 = 1 - 0.68 = 0.32$.

(b) $n = E(Y_1)/p_1 = 10.24/0.32 = 32$ renewals a day.

(c) $p_2 = E(Y_2)/n = 12.16/32 = 0.38$, which leaves $p_3 = 0.30$ for the basic bundle.

(d) $\mathrm{Cov}(Y_1, Y_2) = -np_1p_2 = -32(0.32)(0.38) = -3.8912$. Equivalently, $\mathrm{Cov}(Y_1, Y_2) = -E(Y_1)E(Y_2)/n$.

The ratio V/E identifies q_1 because a binomial count always has variance smaller than its mean. If a reported count had $V(Y_1) > E(Y_1)$, the multinomial model could not fit: the customers would not be independent trials.

Problem 16

seed 20260925

A risk desk classifies each daily move of the Brent crude price as **up** (a rise of more than 1%, probability 0.44), **down** (a fall of more than 1%, probability 0.21) or **quiet** (within 1%, probability 0.35). Days are modelled as independent. Over the next $n = 8$ trading days let Y_1, Y_2, Y_3 be the numbers of up, down and quiet days.

Give each answer as a decimal to at least four significant figures.

(a) The probability of exactly 5 up days and at most one down day, $P(Y_1 = 5,\ Y_2 \le 1) =$ ____________

(b) $P(Y_2 \le 1) =$ ____________

(c) $P(Y_1 = 5 \mid Y_2 \le 1) =$ ____________

(Y_1, Y_2, Y_3) is multinomial with $n = 8$ and $p = (0.44, 0.21, 0.35)$.

(a) The event is the disjoint union of $\{Y_1 = 5, Y_2 = 0, Y_3 = 3\}$ and $\{Y_1 = 5, Y_2 = 1, Y_3 = 2\}$. Definition 5.12 gives

$$p(5,0,3) = \frac{8!}{5!\,0!\,3!}(0.44)^5(0.35)^3 = 56(0.44)^5(0.35)^3 = 0.039596,$$
$$p(5,1,2) = \frac{8!}{5!\,1!\,2!}(0.44)^5(0.21)(0.35)^2 = 168(0.44)^5(0.21)(0.35)^2 = 0.071273.$$

Adding gives $0.039596 + 0.071273 = 0.110870$.

(b) Y_2 is binomial(8, 0.21), so $P(Y_2 \le 1) = (0.79)^8 + 8(0.21)(0.79)^7 = 0.474337$.

(c) $\{Y_1 = 5, Y_2 \le 1\}$ is a subset of $\{Y_2 \le 1\}$, so Definition 2.9 gives $0.110870/0.474337 = 0.233737$.

Problem 17

seed 20260925

An FX dealer at a Baku bank answers $n = 56$ requests for quotes from corporate clients on a typical day. Each request, independently, ends in one of four ways:

	small ticket	mid ticket	large informed ticket	quote not taken
probability	0.44	0.17	0.05	0.34
dealer's P\&L (AZN)	$+54$	$+141$	-310	0

Let Y_1, Y_2, Y_3, Y_4 be the day's counts, so the day's P&L is $P = 54Y_1 + 141Y_2 - 310Y_3$.

Give each answer to at least four significant figures.

(a) $E(P) =$ ____________ AZN

(b) In Theorem 5.12, the variance of P is the sum of the three $a_i^2 V(Y_i)$ terms plus the covariance terms $2\sum_{i<j} a_i a_j \mathrm{Cov}(Y_i, Y_j)$. The covariance terms add up to ____________ AZN^2

(c) $V(P) =$ ____________ AZN^2

(d) The standard deviation of the day's P&L is ____________ AZN

$(Y_1, \dots, Y_4)$ is multinomial with $n = 56$ and the probabilities in the table. $P = a_1Y_1 + a_2Y_2 + a_3Y_3$ with $a_1 = 54$, $a_2 = 141$, $a_3 = -310$; Y_4 has coefficient 0.

(a) $E(P) = 56\,(54(0.44) + 141(0.17) - 310(0.05)) = 1804.88$ AZN.

(b) By Theorem 5.13, $\mathrm{Cov}(Y_1, Y_2) = -np_1p_2 = -4.1888$, $\mathrm{Cov}(Y_1, Y_3) = -1.2320$ and $\mathrm{Cov}(Y_2, Y_3) = -0.4760$. With the signs of the coefficients,

$$2a_1a_2\mathrm{Cov}(Y_1,Y_2) + 2a_1a_3\mathrm{Cov}(Y_1,Y_3) + 2a_2a_3\mathrm{Cov}(Y_2,Y_3) = 15228(-4.1888) - 33480(-1.2320) - 87420(-0.4760) = 19072.23.$$

The two terms involving the loss cell have a negative coefficient times a negative covariance, so they are positive. The contribution here is positive, so V(P) is larger than the sum of the three variance terms. A day with many large informed tickets is, for fixed n, also a day with fewer profitable tickets, so losses and missing profits pile up together.

(c) $V(Y_1) = 13.7984$, $V(Y_2) = 7.9016$ and $V(Y_3) = 2.6600$ (Theorem 5.13), so

$$V(P) = 2916(13.7984) + 19881(7.9016) + 96100(2.6600) + (19072.23) = 452953.84 + (19072.23) = 472026.08.$$

Check. P is also the sum of n independent per-request P&Ls X, so $V(P) = nV(X)$. Here $E(X) = 32.2300$, $E(X^2) = 9467.8100$ and $V(X) = 8429.0371$, so $nV(X) = 472026.08$, in agreement.

(d) $\sigma_P = \sqrt{472026.08} = 687.04$ AZN.

Problem 18

seed 20260925

A bank holds $n = 11$ corporate loans of 1 million AZN each, all to firms rated BBB today. From the rating agency's one-year transition matrix, each firm independently ends the year in one of four states. The bank's year-end charge per loan (capital held, or the loss if the firm defaults) depends on that state:

year-end state	A	BBB	BB	default
probability	0.08	0.84	0.06	0.02
charge (thousand AZN)	4	10	19	570

Let Y_1, Y_2, Y_3, Y_4 be the numbers of firms in the four states and $K = 4Y_1 + 10Y_2 + 19Y_3 + 570Y_4$ the total charge, in thousand AZN.

Give each answer to at least four significant figures.

(a) $E(K) =$ ______________

(b) $V(K) =$ ______________

(c) The standard deviation of K is ______________

(d) The probability that no firm defaults and at most one is downgraded to BB, $P(Y_4 = 0,\ Y_3 \leq 1) =$ ______________

(Y_1, Y_2, Y_3, Y_4) is multinomial with $n = 11$ and the probabilities in the table.

(a) Theorem 5.12 with $E(Y_i) = np_i$: $E(K) = 11\,(4(0.08) + 10(0.84) + 19(0.06) + 570(0.02)) = 233.8600$.

(b) You can use all four counts and the six covariances. It is shorter to remove the largest cell first. Since $Y_2 = 11 - Y_1 - Y_3 - Y_4$,

$$K = 110 + (-6)Y_1 + (9)Y_3 + (560)Y_4,$$

and the constant does not affect the variance. Theorem 5.13 gives $V(Y_1) = 0.8096$, $V(Y_3) = 0.6204$, $V(Y_4) = 0.2156$, $\text{Cov}(Y_1, Y_3) = -0.0528$, $\text{Cov}(Y_1, Y_4) = -0.0176$ and $\text{Cov}(Y_3, Y_4) = -0.0132$. Then Theorem 5.12 gives

$V(K) = 36(0.8096) + 81(0.6204) + 313600(0.2156) + (-108)(-0.0528) + (-6720)(-0.0176) + (10080)(-0.0132) = 67682.4764.$

(The same number, 67682.4764, comes from $nV(X)$, where X is one loan's charge.) Almost all of it comes from the default cell. A single default costs 570, which dwarfs the capital differences between ratings.

(c) $\sqrt{67682.4764} = 260.1586$ thousand AZN.

(d) Group the cells as $\{A, BBB\}$ (probability 0.92), $\{BB\}$ and $\{\text{default}\}$. The grouped counts are multinomial. With no default the event is $Y_3 = 0$ or $Y_3 = 1$, with every other firm in $\{A, BBB\}$:

$$P = \frac{11!}{11!\,0!\,0!}(0.92)^{11} + \frac{11!}{10!\,1!\,0!}(0.92)^{10}(0.06) = 0.399637 + 0.286696 = 0.686334.$$

Problem 19

seed 20260925

A junior trader on a Baku bank's treasury desk opens and closes $n = 6$ positions in a week. Each round trip, independently, closes at a profit (probability 0.49), at a loss (probability 0.28) or flat (probability 0.23). Let Y_1, Y_2, Y_3 be the numbers of winning, losing and flat trades. The desk head watches the net win count $D = Y_1 - Y_2$.

Give each answer as a decimal to at least four significant figures.

(a) The probability that the week has as many losing as winning trades, $P(Y_1 = Y_2) =$ ______________

(b) $E(Y_1 Y_2) =$ ______________

(c) $V(D) =$ ______________

(d) $E(D^2) =$ ______________

(Y_1, Y_2, Y_3) is multinomial with $n = 6$ and $p = (0.49, 0.28, 0.23)$.

(a) $\{Y_1 = Y_2\}$ is the union of the disjoint events $\{Y_1 = j, Y_2 = j, Y_3 = 6 - 2j\}$ for $j = 0, 1, \ldots, 3$. Definition 5.12 gives

$$P(Y_1 = Y_2) = \sum_{j=0}^{3} \frac{6!}{j!\, j!\, (6-2j)!}(0.49)^j (0.28)^j (0.23)^{6-2j} = 0.000148 + 0.011518 + 0.089620 + 0.051653 = 0.152939.$$

(b) Rearrange the definition of covariance: $E(Y_1Y_2) = \mathrm{Cov}(Y_1, Y_2) + E(Y_1)E(Y_2)$. Theorem 5.13 gives

$$E(Y_1Y_2) = -np_1p_2 + (np_1)(np_2) = n(n-1)p_1p_2 = -0.8232 + 4.9392 = 4.1160.$$

Summing $y_1y_2\, p(y_1, y_2, y_3)$ over the whole table would give the same number with far more work.

(c) Theorem 5.12 with coefficients 1 and -1:

$$V(D) = np_1q_1 + np_2q_2 + 2np_1p_2 = n\left\{(p_1 + p_2) - (p_1 - p_2)^2\right\} = 6\left\{0.77 - (0.21)^2\right\} = 6(0.7259) = 4.3554.$$

(d) $E(D) = n(p_1 - p_2) = 1.26$, so $E(D^2) = V(D) + (E(D))^2 = 4.3554 + (1.26)^2 = 5.9430$.

Problem 20

seed 20260925

A quant at an Azerbaijani asset manager prices a digital option on a stock index by simulation. In each of $n = 80$ independent scenarios the index finishes in one of three regions:

- above the strike, so the option pays off (probability $p_1 = 0.28$);
- below its starting level (probability $p_2 = 0.40$);
- between the start and the strike (probability $p_3 = 0.32$).

Let Y_1, Y_2, Y_3 be the numbers of scenarios in each region. The quant needs np_1. She knows p_2 exactly from a closed-form formula, so she can use Y_2 as a **control variate**:

$$U = Y_1 + c\,(Y_2 - np_2),$$

which has $E(U) = np_1$ for every constant c.

Give each answer to at least four significant figures.

(a) $\mathrm{Cov}(Y_1, Y_2) =$ ______________

(b) The value of c that minimises $V(U)$ is ______________

(c) The minimum value of $V(U)$ is ______________

(d) The fraction by which this reduces the variance, compared with using Y_1 alone, $1 - V_{\min}(U)/V(Y_1)$, is ______________

(Y_1, Y_2, Y_3) is multinomial with $n = 80$ and $p = (0.28, 0.40, 0.32)$. $E(U) = np_1 + c(np_2 - np_2) = np_1$ by Theorem 5.13, so U is unbiased whatever c is.

(a) $\mathrm{Cov}(Y_1, Y_2) = -np_1p_2 = -80(0.28)(0.40) = -8.9600$.

(b) The constant $-cnp_2$ does not affect the variance. Theorem 5.12 gives

$$V(U) = V(Y_1) + c^2V(Y_2) + 2c\,\mathrm{Cov}(Y_1, Y_2),$$

a quadratic in c. Setting $dV/dc = 2cV(Y_2) + 2\mathrm{Cov}(Y_1, Y_2) = 0$ gives

$$c^* = -\frac{\mathrm{Cov}(Y_1, Y_2)}{V(Y_2)} = \frac{np_1p_2}{np_2q_2} = \frac{p_1}{1 - p_2} = \frac{0.28}{0.60} = 0.466667.$$

The coefficient of c^2 is positive, so this is a minimum. The optimal c is positive: the covariance is negative, so an unusually low Y_2 signals that Y_1 is probably too high, and the correction pulls it back.

(c) Substitute c^*: $V_{\min}(U) = V(Y_1) - \mathrm{Cov}(Y_1, Y_2)^2/V(Y_2)$. Here $V(Y_1) = 80(0.28)(0.72) = 16.1280$ and $V(Y_2) = 80(0.40)(0.60) = 19.2000$, so

$$V_{\min}(U) = 16.1280 - \frac{(-8.9600)^2}{19.2000} = 11.9467.$$

(d) $1 - 11.9467/16.1280 = 0.259259$. This equals ρ^2, where $\rho = -0.509175$ is the correlation of Y_1 and Y_2. It also equals $p_1 p_2/(q_1 q_2)$, so it does not depend on n.

The choice of c matters. With $c = 1$, $V(U) = V(Y_1) + V(Y_2) + 2\mathrm{Cov}(Y_1, Y_2) = 17.4080$, compared with $V(Y_1) = 16.1280$. So a careless c = 1 is worse than using no control at all.

Answer key

Problem	Answers
1	(a) 0.0945504, (b) 3.08, (c) 1.4973
2	(a) 4.29, (b) 2.7104, (c) 1.24583, (d) -0.7293
3	(a) -0.1134, (b) -0.9702, (c) -0.126886, (d) Upgrades and ... staying BBB
4	(a) 0.38742, (b) 0.120869, (c) 0.225159, (d) 6.66
5	(a) 0.08598, (b) 0.308715, (c) 1.84
6	(a) 0.0212497, (b) 0.404567
7	(a) -2.4024, (b) 4.76, (c) 16.2708, (d) 4.03371
8	(a) 178.88, (b) 1924.67, (c) 43.871
9	(a) 0.2, (b) 0.367002, (c) 1.4, (d) 1.12
10	(a) 0.252089, (b) 0.0154646, (c) 0.105261
11	(a) 0.540137, (b) 0.213123
12	(a) 10.08, (b) -0.3136, (c) 8.2656, (d) -0.0823586
13	(a) 0.21, (b) 0.00235792, (c) 0.0485584
14	(a) 0.0611821, (b) 1.11236
15	(a) 0.32, (b) 32, (c) 0.38, (d) -3.8912
16	(a) 0.11087, (b) 0.474337, (c) 0.233737
17	(a) 1804.88, (b) 19072.2, (c) 472026, (d) 687.042
18	(a) 233.86, (b) 67682.5, (c) 260.159, (d) 0.686334
19	(a) 0.152939, (b) 4.116, (c) 4.3554, (d) 5.943
20	(a) -8.96, (b) 0.466667, (c) 11.9467, (d) 0.259259