Week 12, Problem Set 2 - worked solutions

Wackerly 5.3 - Marginal and Conditional Probability Distributions

> 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 Baku household budget survey classifies each household by its income band Y_1 and its monthly consumption band Y_2, each coded $1 =$ low, $2 =$ middle, $3 =$ high. For a randomly chosen household the joint probability function $p(y_1, y_2)$ is

$y_2 \backslash y_1$	1	2	3
1	0.07	0.05	0.04
2	0.04	0.46	0.07
3	0.04	0.08	0.15

(rows are values of y_2, columns are values of y_1).

(a) $p_1(1) = P(Y_1 = 1) =$ ____________

(b) $p_1(3) = P(Y_1 = 3) =$ ____________

(c) $p_2(1) = P(Y_2 = 1) =$ ____________

(d) $p_2(3) = P(Y_2 = 3) =$ ____________

(e) $P(Y_1 \geq 2, Y_2 \geq 2) =$ ____________

By Definition 5.4(a), $p_1(y_1) = \sum_{\text{all } y_2} p(y_1, y_2)$: to get the marginal of Y_1 we add down each **column** of the table, and to get the marginal of Y_2 we add across each **row**.

(a) $p_1(1) = p(1,1) + p(1,2) + p(1,3) = 0.07 + 0.04 + 0.04 = 0.15$.

(b) $p_1(3) = 0.04 + 0.07 + 0.15 = 0.26$.

For completeness $p_1(2) = 0.05 + 0.46 + 0.08 = 0.59$, and $0.15 + 0.59 + 0.26 = 1$, as a probability function must.

(c) $p_2(1) = p(1,1) + p(2,1) + p(3,1) = 0.07 + 0.05 + 0.04 = 0.16$.

(d) $p_2(3) = 0.04 + 0.08 + 0.15 = 0.27$. (Also $p_2(2) = 0.57$, and the three add to 1.)

(e) This is a joint event, so we add the four cells with $y_1 \geq 2$ and $y_2 \geq 2$: $p(2,2) + p(2,3) + p(3,2) + p(3,3) = 0.46 + 0.08 + 0.07 + 0.15 = 0.76$.

The marginals are literally the totals written in the margins of the table, which is where the name comes from. Note what they throw away: p_1 tells a lender how incomes are spread, but nothing about how consumption moves with income. That information lives only in the joint table.

Problem 2

seed 20260925

A bank's consumer-credit book is classified by loan size Y_1 ($1 =$ small, $2 =$ medium, $3 =$ large) and by outcome Y_2 (1 if the loan defaulted within two years, 0 if it did not). For a loan drawn at random from the book, the joint probability function $p(y_1, y_2)$ is

$y_2 \backslash y_1$	1	2	3
0	0.38	0.36	0.17
1	0.02	0.05	0.02

Give every probability to at least four significant figures.

(a) The marginal default probability $p_2(1) = P(Y_2 = 1) =$ ____________

(b) The default rate among large loans, $P(Y_2 = 1 \mid Y_1 = 3) =$ ____________

(c) The share of defaulted loans that were large, $P(Y_1 = 3 \mid Y_2 = 1) =$ ____________

(d) $P(Y_2 = 0 \mid Y_1 = 1) =$ ____________

(a) Definition 5.4(a): sum the row $y_2 = 1$ over all y_1, $p_2(1) = 0.02 + 0.05 + 0.02 = 0.09$.

The column totals give the marginal of loan size: $p_1(1) = 0.40$, $p_1(2) = 0.41$, $p_1(3) = 0.19$.

(b) Definition 5.5 with the roles of the variables swapped, $p(y_2 \mid y_1) = p(y_1, y_2)/p_1(y_1)$, defined because $p_1(3) > 0$:

$$P(Y_2 = 1 \mid Y_1 = 3) = \frac{p(3,1)}{p_1(3)} = \frac{0.02}{0.19} = 0.1053.$$

(c) Now condition on the outcome:

$$P(Y_1 = 3 \mid Y_2 = 1) = \frac{p(3,1)}{p_2(1)} = \frac{0.02}{0.09} = 0.2222.$$

Parts (b) and (c) share a numerator and differ only in which margin they divide by. (b) is a **default rate** (what a risk officer prices); (c) is a **composition of the loss pool** (what a recovery team sees). Confusing them is a classic reporting error.

(d) $P(Y_2 = 0 \mid Y_1 = 1) = 0.38/0.40 = 0.9500$. As a check, $P(Y_2 = 1 \mid Y_1 = 1) = 0.02/0.40 = 0.0500$, and the two conditional probabilities add to 1: for fixed y_1, $p(y_2 \mid y_1)$ is itself a probability function.

Problem 3

seed 20260925

A bank monitors two revolving credit lines of the same corporate client. Let Y_1 and Y_2 be the proportions of the two limits that are drawn at the end of a randomly chosen month. The bank's model for their joint density is

$$f(y_1, y_2) = \begin{cases} 0.60\,y_1 + 1.40\,y_2, & 0 \le y_1 \le 1,\ 0 \le y_2 \le 1, \\ 0, & \text{elsewhere.} \end{cases}$$

(a) For $0 \le y_1 \le 1$, the marginal density of Y_1 is $f_1(y_1) =$ ______________

(a formula in y_1; type it in the form $3 * y_1 + 0.2$)

(b) $P(Y_1 \le 0.85) =$ ______________

(c) $P(Y_2 > 0.40) =$ ______________

Give probabilities to at least four significant figures.

(a) Definition 5.4(b): integrate y_2 out over its whole range. For $0 \le y_1 \le 1$ the density is non-zero only for $0 \le y_2 \le 1$, so

$$f_1(y_1) = \int_0^1 (0.60\,y_1 + 1.40\,y_2)\,dy_2 = 0.60\,y_1 + 1.40\left[\frac{y_2^2}{2}\right]_0^1 = 0.60\,y_1 + 0.70,$$

and $f_1(y_1) = 0$ outside $[0, 1]$. It integrates to $0.60/2 + 0.70 = 0.30 + 0.70 = 1$, as it must.

(b) Once we have the marginal, a statement about Y_1 alone is a one-variable integral:

$$P(Y_1 \le 0.85) = \int_0^{0.85} (0.60\,y_1 + 0.70)\,dy_1 = 0.30(0.85)^2 + 0.70(0.85) = 0.8117.$$

(c) By the same argument with the roles swapped, $f_2(y_2) = \int_0^1 (0.60\,y_1 + 1.40\,y_2)\,dy_1 = 1.40\,y_2 + 0.30$ for $0 \le y_2 \le 1$. Then

$$P(Y_2 \le 0.40) = 0.70(0.40)^2 + 0.30(0.40) = 0.2320,$$

so $P(Y_2 > 0.40) = 1 - 0.2320 = 0.7680$.

The two marginals are different straight-line densities even though the joint density treats both lines symmetrically in form: the larger coefficient tilts its own variable's marginal towards full utilisation.

Problem 4

seed 20260925

A mobile operator in Azerbaijan studies postpaid billing. For a randomly chosen postpaid account let Y_2 be the segment ($1 =$ consumer, $2 =$ business) and Y_1 the number of bills paid late in the last quarter. The joint probability function $p(y_1, y_2)$ is

$y_2 \backslash y_1$	0	1	2	3
1	0.28	0.15	0.05	0.04
2	0.35	0.05	0.05	0.03

Give every probability to at least four significant figures.

(a) The conditional probability function of Y_1 given $Y_2 = 2$:

$p(0 \mid 2) =$ ______________

$p(1 \mid 2) =$ ______________

$p(2 \mid 2) =$ ______________

(b) The probability that a consumer account paid at least one bill late, $P(Y_1 \geq 1 \mid Y_2 = 1) =$ ______________

(a) By Definition 5.5, $p(y_1 \mid y_2) = p(y_1, y_2)/p_2(y_2)$, provided $p_2(y_2) > 0$. First the marginal: $p_2(2) = 0.35 + 0.05 + 0.05 + 0.03 = 0.48$. Then we divide the business row by it:

$$p(0 \mid 2) = \frac{0.35}{0.48} = 0.7292, \quad p(1 \mid 2) = \frac{0.05}{0.48} = 0.1042, \quad p(2 \mid 2) = \frac{0.05}{0.48} = 0.1042,$$

and $p(3 \mid 2) = 0.03/0.48 = 0.0625$. The four add to 1: conditioning on $Y_2 = 2$ keeps only the business row and rescales it into a probability function.

(b) $p_2(1) = 0.28 + 0.15 + 0.05 + 0.04 = 0.52$, and

$$P(Y_1 \geq 1 \mid Y_2 = 1) = \frac{p(1,1) + p(2,1) + p(3,1)}{p_2(1)} = \frac{0.24}{0.52} = 0.4615.$$

(Equivalently $1 - p(0 \mid 1)$.) A collections team sizing its call list for each segment needs exactly these conditional figures, not the joint cells.

Problem 5

seed 20260925

An underground gas storage facility near Baku has working capacity 70 million cubic metres. Let Y_2 be the volume in storage at the start of a winter month and Y_1 the volume withdrawn during the month (both in million cubic metres). The facility is not refilled during the month, so $Y_1 \leq Y_2$. The operator models (Y_1, Y_2) as uniformly distributed over the triangle $0 \leq y_1 \leq y_2 \leq 70$:

$$f(y_1, y_2) = \begin{cases} k, & 0 \leq y_1 \leq y_2 \leq 70, \\ 0, & \text{elsewhere.} \end{cases}$$

Give every answer to at least four significant figures (decimals, not fractions of large numbers, are fine).

(a) The constant $k =$ ______________

(b) The marginal density of Y_2 at $y_2 = 56$: $f_2(56) =$ ______________

(c) $P(Y_1 \leq 13 \mid Y_2 = 56) =$ ______________

(d) $P(Y_1 \leq 13 \mid Y_2 = 70) =$ ______________

(a) The triangle has area $\frac{1}{2} \cdot 70 \cdot 70 = 4900/2$, so a constant density must be $k = 2/70^2 = 2/4900 = 0.00040816$.

(b) Definition 5.4(b), integrating y_1 out. For fixed $0 \leq y_2 \leq 70$ the density is non-zero only for $0 \leq y_1 \leq y_2$, so the upper limit depends on y_2:

$$f_2(y_2) = \int_0^{y_2} k\, dy_1 = k\, y_2 = \frac{2y_2}{4900}, \qquad 0 \leq y_2 \leq 70.$$

So $f_2(56) = 2(56)/4900 = 0.0228571$.

(c) Definition 5.7: for $0 < y_2 \leq 70$,

$$f(y_1 \mid y_2) = \frac{f(y_1, y_2)}{f_2(y_2)} = \frac{k}{k\, y_2} = \frac{1}{y_2}, \qquad 0 \leq y_1 \leq y_2.$$

Given the storage level, the withdrawal is uniform on $(0, y_2)$. Hence

$$P(Y_1 \leq 13 \mid Y_2 = 56) = \int_0^{13} \frac{1}{56}\, dy_1 = \frac{13}{56} = 0.2321.$$

(d) Same conditional density with $y_2 = 70$: $13/70 = 0.1857$.

The conditional probability of a small withdrawal changes appreciably with the starting level: a fuller facility makes a small draw less likely. That dependence is invisible in either marginal alone.

Problem 6

seed 20260925

On the Baku Stock Exchange, let Y_1 and Y_2 be the times (in minutes, from 11:00) until the next trade in stock A and in stock B. A market-microstructure analyst uses the joint density

$$f(y_1, y_2) = \begin{cases} \dfrac{1}{25.50}\, e^{-y_1/3.0 \,-\, y_2/8.5}, & y_1 > 0,\ y_2 > 0, \\ 0, & \text{elsewhere.} \end{cases}$$

(Note $25.50 = 3.0 \times 8.5$.) Give probabilities to at least four significant figures.

(a) For $y_1 > 0$, $f_1(y_1) =$ ______________

(a formula in y_1; use $e^(\ldots)$ or $exp(\ldots)$)

(b) $P(2.0 < Y_1 < 6.5) =$ ______________

(c) $P(2.0 < Y_1 < 6.5 \mid Y_2 = 5) =$ ______________

(d) $P(Y_2 > 11) =$ ______________

(a) Definition 5.4(b): for $y_1 > 0$,

$$f_1(y_1) = \int_0^\infty \frac{1}{3.0 \cdot 8.5} e^{-y_1/3.0} e^{-y_2/8.5}\, dy_2 = \frac{1}{3.0} e^{-y_1/3.0} \int_0^\infty \frac{1}{8.5} e^{-y_2/8.5}\, dy_2 = \frac{1}{3.0} e^{-y_1/3.0},$$

because the last integral is the total probability of an exponential density. So Y_1 is exponential with $\beta = 3.0$, and likewise $f_2(y_2) = \frac{1}{8.5} e^{-y_2/8.5}$, $y_2 > 0$.

(b) $P(2.0 < Y_1 < 6.5) = e^{-2.0/3.0} - e^{-6.5/3.0} = 0.5134 - 0.1146 = 0.3989$.

(c) Definition 5.7: for any $y_2 > 0$,

$$f(y_1 \mid y_2) = \frac{f(y_1, y_2)}{f_2(y_2)} = \frac{\frac{1}{3.0 \cdot 8.5} e^{-y_1/3.0} e^{-y_2/8.5}}{\frac{1}{8.5} e^{-y_2/8.5}} = \frac{1}{3.0} e^{-y_1/3.0} = f_1(y_1).$$

The conditional density does not depend on y_2 at all, so the conditional probability equals the marginal one from (b): 0.3989, whatever value $y_2 = 5$ takes. Learning when stock B next trades tells the analyst nothing about stock A. (Section 5.4 names this property.)

(d) $P(Y_2 > 11) = \int_{11}^\infty \frac{1}{8.5} e^{-y_2/8.5}\, dy_2 = e^{-11/8.5} = 0.2741$.

Problem 7

seed 20260925

A consumer-lending unit uses a household survey in which Y_1 is the household's income band and Y_2 its consumption band, each coded $1 =$ low, $2 =$ middle, $3 =$ high. The joint probability function $p(y_1, y_2)$ is

$y_2 \backslash y_1$	1	2	3
1	0.09	0.07	0.03
2	0.05	0.41	0.08
3	0.03	0.10	0.14

Give every probability to at least four significant figures.

(a) The conditional probability function of consumption for high-income households:

$P(Y_2 = 1 \mid Y_1 = 3) =$ ______________

$P(Y_2 = 2 \mid Y_1 = 3) =$ ______________

$P(Y_2 = 3 \mid Y_1 = 3) =$ ______________

(b) Among high-consumption households, the probability of high income, $P(Y_1 = 3 \mid Y_2 = 3) =$ ______________

(c) $P(Y_2 = 3 \mid Y_1 = 1) =$ ______________

(a) Definition 5.5 (with Y_1 as the conditioning variable): $p(y_2 \mid y_1) = p(y_1, y_2)/p_1(y_1)$. The column $y_1 = 3$ has total $p_1(3) = 0.03 + 0.08 + 0.14 = 0.25$, so

$$p(1 \mid 3) = \frac{0.03}{0.25} = 0.1200, \quad p(2 \mid 3) = \frac{0.08}{0.25} = 0.3200, \quad p(3 \mid 3) = \frac{0.14}{0.25} = 0.5600,$$

which add to 1.

(b) Now condition on consumption. The row $y_2 = 3$ has total $p_2(3) = 0.03 + 0.10 + 0.14 = 0.27$, so

$$P(Y_1 = 3 \mid Y_2 = 3) = \frac{p(3,3)}{p_2(3)} = \frac{0.14}{0.27} = 0.5185.$$

(c) $p_1(1) = 0.09 + 0.05 + 0.03 = 0.17$, so $P(Y_2 = 3 \mid Y_1 = 1) = 0.03/0.17 = 0.1765$.

Compare with the marginal $p_2(3) = 0.27$. Knowing that a household has high income multiplies the probability of high consumption by about $0.5600/0.27 \approx 2.07$; knowing it has low income pushes the probability down to 0.1765. The conditional distributions differ from column to column, which is exactly what the marginals cannot show.

Problem 8

seed 20260925

A Baku commercial bank's loan book, by size band Y_1 ($1 =$ micro, $2 =$ SME, $3 =$ corporate) and status Y_2 ($0 =$ current, $1 =$ 30-89 days past due, $2 =$ defaulted), contains the following **numbers of loans**:

$y_2 \backslash y_1$	1	2	3
0	480	470	180
1	67	25	11
2	28	32	2

A supervisor draws one loan at random from the whole book, so that $p(y_1, y_2)$ is the count in cell (y_1, y_2) divided by the total. Give every probability to at least four significant figures.

(a) $p_1(3) = P(Y_1 = 3) =$ ______________

(b) The default rate of SME loans, $P(Y_2 = 2 \mid Y_1 = 2) =$ ______________

(c) The SME share of defaulted loans, $P(Y_1 = 2 \mid Y_2 = 2) =$ ______________

(d) $P(Y_2 \geq 1 \mid Y_1 = 1) =$ ______________

The total is $N = 1295$ loans. Column totals: micro 575, SME 527, corporate 193; the defaulted row totals 62.

(a) Definition 5.4(a) sums over all y_2: $p_1(3) = 193/1295 = 0.1490$.

(b) Definition 5.5:

$$P(Y_2 = 2 \mid Y_1 = 2) = \frac{p(2,2)}{p_1(2)} = \frac{32/1295}{527/1295} = \frac{32}{527} = 0.0607.$$

The common factor $1/1295$ cancels, so a conditional probability from a count table is just a cell divided by its column total. (In probabilities: $p(2,2) = 0.02471$, $p_1(2) = 0.4069$.)

(c) Condition on the row instead:

$$P(Y_1 = 2 \mid Y_2 = 2) = \frac{p(2,2)}{p_2(2)} = \frac{32}{62} = 0.5161,$$

where $p_2(2) = 62/1295 = 0.04788$.

(d) Given $Y_1 = 1$ we work inside the micro column and add the conditional probabilities of $y_2 = 1$ and $y_2 = 2$:

$$P(Y_2 \geq 1 \mid Y_1 = 1) = \frac{67 + 28}{575} = \frac{95}{575} = 0.1652.$$

Problem 9

seed 20260925

A Central Bank examiner audits a branch whose small-business loan files consist of 6 performing loans, 8 watch-list loans and 4 non-performing loans (18 files in total). She selects 3 files at random, without replacement. Let Y_1 be the number of performing and Y_2 the number of watch-list files in her sample, so that

$$p(y_1, y_2) = \frac{\binom{6}{y_1}\binom{8}{y_2}\binom{4}{3-y_1-y_2}}{\binom{18}{3}}, \qquad y_1, y_2 \geq 0,\ y_1 + y_2 \leq 3.$$

Give every probability to at least four significant figures.

(a) The marginal probability $p_1(3) = P(Y_1 = 3) =$ ______________

(b) $P(Y_1 = 1 \mid Y_2 = 1) =$ ______________

(c) $P(Y_1 = 2 \mid Y_2 = 0) =$ ______________

(d) $p_2(0) = P(Y_2 = 0) =$ ______________

(a) By Definition 5.4(a), $p_1(y_1) = \sum_{y_2} p(y_1, y_2)$. Summing over the ways to split the remaining $3 - y_1$ files between the watch-list and non-performing groups is the same as choosing them from the $8 + 4 = 12$ watch-list and non-performing files together (Vandermonde's identity), so Y_1 is hypergeometric (Section 3.7):

$$p_1(3) = \frac{\binom{6}{3}\binom{12}{0}}{\binom{18}{3}} = \frac{20 \cdot 1}{816} = 0.0245.$$

(b) Definition 5.5. Given $Y_2 = 1$, one watch-list file has been drawn and the other two came from the $6 + 4 = 10$ remaining performing-or-non-performing files. The factor $\binom{8}{1}$ and the denominator $\binom{18}{3}$ cancel between $p(1,1)$ and $p_2(1)$:

$$P(Y_1 = 1 \mid Y_2 = 1) = \frac{p(1,1)}{p_2(1)} = \frac{\binom{6}{1}\binom{4}{1}}{\binom{10}{2}} = \frac{24}{45} = 0.5333.$$

(c) In the same way, given $Y_2 = 0$ all three files came from the 10 others:

$$P(Y_1 = 2 \mid Y_2 = 0) = \frac{\binom{6}{2}\binom{4}{1}}{\binom{10}{3}} = \frac{60}{120} = 0.5000.$$

(d) Y_2 is hypergeometric too, so $p_2(0) = \binom{10}{3}/\binom{18}{3} = 120/816 = 0.1471$.

Conditioning on the number of watch-list files turns the problem into a smaller hypergeometric one on the files that are left, the same structure Example 5.7 exploits.

Problem 10

seed 20260925

A fuel depot restocks its storage tank on Monday mornings. Let Y_1 be the proportion of the tank's capacity stocked on Monday and Y_2 the proportion of capacity sold during the week, so $Y_2 \le Y_1$. The joint density is

$$f(y_1, y_2) = \begin{cases} 3\,y_1, & 0 \le y_2 \le y_1 \le 1, \\ 0, & \text{elsewhere.} \end{cases}$$

(a) For $0 \le y_2 \le 1$, the marginal density of Y_2 is $f_2(y_2) =$ ______________

(a formula in y_2)

(b) Given that the depot stocked 0.90 of capacity, the probability that it sells more than 0.25 of capacity: $P(Y_2 > 0.25 \mid Y_1 = 0.90) =$ ______________

(c) Given that it sold 0.45 of capacity, the probability that it had stocked more than 0.70: $P(Y_1 > 0.70 \mid Y_2 = 0.45) =$ ______________

Give probabilities to at least four significant figures.

(a) Definition 5.4(b). Fix y_2 in $[0,1]$. The density is non-zero only where $y_2 \le y_1 \le 1$, so the **lower** limit of integration is y_2:

$$f_2(y_2) = \int_{y_2}^{1} 3\,y_1\,dy_1 = \frac{3}{2}\left[y_1^2\right]_{y_2}^{1} = \frac{3}{2}(1 - y_2^2).$$

Similarly $f_1(y_1) = \int_0^{y_1} 3\,y_1\,dy_2 = 3\,y_1^2$ for $0 \le y_1 \le 1$.

(b) Definition 5.7: for $0 < y_1 \le 1$,

$$f(y_2 \mid y_1) = \frac{f(y_1, y_2)}{f_1(y_1)} = \frac{3\,y_1}{3\,y_1^2} = \frac{1}{y_1}, \qquad 0 \le y_2 \le y_1.$$

Given the stock, sales are uniform on $(0, y_1)$. So

$$P(Y_2 > 0.25 \mid Y_1 = 0.90) = \int_{0.25}^{0.90} \frac{1}{0.90}\,dy_2 = 1 - \frac{0.25}{0.90} = 0.7222.$$

(c) For $0 \le y_2 < 1$ (where $f_2(y_2) > 0$),

$$f(y_1 \mid y_2) = \frac{3\,y_1}{\frac{3}{2}(1 - y_2^2)} = \frac{2\,y_1}{1 - y_2^2}, \qquad y_2 \le y_1 \le 1.$$

Hence

$$P(Y_1 > 0.70 \mid Y_2 = 0.45) = \int_{0.70}^{1} \frac{2\,y_1}{1 - 0.45^2}\,dy_1 = \frac{1 - 0.70^2}{1 - 0.45^2} = \frac{1 - 0.4900}{1 - 0.2025} = 0.6395.$$

The two conditional densities look nothing alike: one is flat, the other rises towards a full tank. The support is a triangle, so each conditional lives on an interval whose end depends on the value conditioned on.

Problem 11

seed 20260925

A ministry tracks capital projects over a fiscal year. For a randomly chosen project let Y_2 be the proportion of its budget **committed** (contracts signed) by the end of Q3 and Y_1 the proportion actually **disbursed**; nothing is disbursed before it is committed, so $Y_1 \le Y_2$. The joint density is

$$f(y_1, y_2) = \begin{cases} 6(1 - y_2), & 0 \le y_1 \le y_2 \le 1, \\ 0, & \text{elsewhere.} \end{cases}$$

(a) For $0 \le y_1 \le 1$, $f_1(y_1) =$ ______________ (a formula in y_1)

(b) $P(Y_2 \le 0.50 \mid Y_1 \le 0.40) =$ ______________

(c) $P(Y_2 \ge 0.60 \mid Y_1 = 0.35) =$ ______________

Give probabilities to at least four significant figures.

(a) Definition 5.4(b). For fixed y_1, y_2 runs from y_1 to 1:

$$f_1(y_1) = \int_{y_1}^{1} 6(1-y_2)\,dy_2 = \left[-3(1-y_2)^2\right]_{y_1}^{1} = 3(1-y_1)^2, \qquad 0 \le y_1 \le 1.$$

(b) This conditions on an **event** of positive probability, so it is the Chapter 2 definition, $P(A \mid B) = P(A \cap B)/P(B)$. The denominator is

$$P(Y_1 \le 0.40) = \int_0^{0.40} 3(1-y_1)^2\,dy_1 = 1-(1-0.40)^3 = 1-0.60^3 = 0.78400.$$

For the numerator, since $0.50 \ge 0.40$, every $y_1 \le 0.40$ allows y_2 from y_1 up to 0.50:

$$P(Y_1 \le 0.40, Y_2 \le 0.50) = \int_0^{0.40}\int_{y_1}^{0.50} 6(1-y_2)\,dy_2\,dy_1 = \int_0^{0.40} 6\left[\left(0.50-\tfrac{0.50^2}{2}\right)-\left(y_1-\tfrac{y_1^2}{2}\right)\right]dy_1$$

$$= 6\left[0.37500\,(0.40) - \frac{0.40^2}{2} + \frac{0.40^3}{6}\right] = 0.48400.$$

So $P(Y_2 \le 0.50 \mid Y_1 \le 0.40) = 0.48400/0.78400 = 0.6173$.

(c) This conditions on a **value** of a continuous variable, a zero-probability event, so we need Definition 5.7. For $0 \le y_1 < 1$,

$$f(y_2 \mid y_1) = \frac{6(1-y_2)}{3(1-y_1)^2} = \frac{2(1-y_2)}{(1-y_1)^2}, \qquad y_1 \le y_2 \le 1.$$

Then

$$P(Y_2 \ge 0.60 \mid Y_1 = 0.35) = \int_{0.60}^{1} \frac{2(1-y_2)}{0.65^2}\,dy_2 = \frac{(1-0.60)^2}{0.65^2} = \left(\frac{0.40}{0.65}\right)^2 = 0.3787.$$

Problem 12

seed 20260925

A bank's FX department runs two market-making desks. Over a randomly chosen trading day let Y_1 and Y_2 be the proportions of the session during which desk I and desk II are actively quoting two-way prices. Their joint density is

$$f(y_1, y_2) = \begin{cases} y_1 + y_2, & 0 \le y_1 \le 1,\ 0 \le y_2 \le 1, \\ 0, & \text{elsewhere.} \end{cases}$$

Give every probability to at least four significant figures.

(a) On a day when desk II quotes for exactly 0.85 of the session, the probability that desk I quotes for more than 0.30 of it: $P(Y_1 > 0.30 \mid Y_2 = 0.85) =$ ____________

(b) $P(Y_1 \ge 0.45 \mid Y_2 \ge 0.75) =$ ____________

(c) Unconditionally, $P(Y_1 > 0.30) =$ ____________

First the marginals (Definition 5.4(b)): for $0 \le y_2 \le 1$, $f_2(y_2) = \int_0^1 (y_1+y_2)\,dy_1 = y_2 + \frac{1}{2}$, and by symmetry $f_1(y_1) = y_1 + \frac{1}{2}$ for $0 \le y_1 \le 1$.

(a) Definition 5.7:

$$f(y_1 \mid y_2) = \frac{y_1 + y_2}{y_2 + \frac{1}{2}}, \qquad 0 \le y_1 \le 1.$$

With $y_2 = 0.85$,

$$P(Y_1 > 0.30 \mid Y_2 = 0.85) = \frac{1}{1.35}\int_{0.30}^{1} (y_1 + 0.85)\,dy_1 = \frac{\frac{1-0.30^2}{2} + 0.85(1-0.30)}{1.35} = \frac{0.45500 + 0.59500}{1.35} = 0.7778.$$

(b) Here we condition on an event of positive probability:

$$P(Y_1 \ge 0.45, Y_2 \ge 0.75) = \int_{0.75}^{1}\int_{0.45}^{1} (y_1+y_2)\,dy_1\,dy_2 = (1-0.75)\frac{1-0.45^2}{2} + (1-0.45)\frac{1-0.75^2}{2} = 0.22000,$$

$$P(Y_2 \ge 0.75) = \int_{0.75}^{1} (y_2 + \tfrac{1}{2})\,dy_2 = \frac{1-0.75^2}{2} + \frac{1-0.75}{2} = 0.34375,$$

so the conditional probability is $0.22000/0.34375 = 0.6400$.

(c) $P(Y_1 > 0.30) = \int_{0.30}^{1} (y_1 + \frac{1}{2})\,dy_1 = \frac{1-0.30^2}{2} + \frac{1-0.30}{2} = 0.8050$.

Compare (a) with (c). Writing $A = (1-0.30^2)/2$ and $B = 1-0.30$, the difference is

$$\frac{A + y_2 B}{y_2 + \frac{1}{2}} - \left(A + \frac{B}{2}\right) = \frac{(y_2 - \frac{1}{2})(\frac{B}{2} - A)}{y_2 + \frac{1}{2}},$$

and $B/2 - A = -0.30(1-0.30)/2 < 0$. So the conditional probability exceeds the marginal one exactly when $y_2 < \frac{1}{2}$. That may be the opposite of what you guessed: dividing $y_1 + y_2$ by its total makes the conditional density **steeper** in y_1 when y_2 is small. Under this model a quiet day for desk II tilts desk I towards a busy one, so the two desks behave a little like substitutes. The conditional distribution, not intuition about the joint formula, is what settles it.

Problem 13

seed 20260925

For a consumer-loan application at a Baku bank let Y_1 be the total time (in working days) from submission to disbursement and Y_2 the time the file waits in the underwriting queue before anyone opens it, so $Y_2 \le Y_1$. The bank's operations team fits

$$f(y_1, y_2) = \begin{cases} \dfrac{1}{9.00}\, e^{-y_1/3.0}, & 0 \le y_2 \le y_1 < \infty, \\ 0, & \text{elsewhere,} \end{cases}$$

where $9.00 = 3.0^2$. Give probabilities to at least four significant figures.

(a) For $y_1 > 0$, $f_1(y_1) =$ ____________ (a formula in y_1; use $e^(\dots)$ or $exp(\dots)$)

(b) $P(Y_2 > 6) =$ ____________

(c) An application took 8 days in total. The probability that it waited in the queue less than 3.5 days: $P(Y_2 < 3.5 \mid Y_1 = 8) =$ ____________

(d) An application waited 4 days in the queue. The probability that it is still not disbursed 6 days after leaving the queue: $P(Y_1 > 10 \mid Y_2 = 4) =$

____

(a) Definition 5.4(b); for fixed $y_1 > 0$, y_2 runs over $[0, y_1]$:

$$f_1(y_1) = \int_0^{y_1} \frac{1}{9.00} e^{-y_1/3.0}\, dy_2 = \frac{y_1\, e^{-y_1/3.0}}{3.0^2}, \qquad y_1 > 0,$$

a gamma density with $\alpha = 2$, $\beta = 3.0$ (Section 4.6).

(b) For fixed $y_2 \ge 0$, y_1 runs over $[y_2, \infty)$:

$$f_2(y_2) = \int_{y_2}^{\infty} \frac{1}{9.00} e^{-y_1/3.0}\, dy_1 = \frac{1}{3.0} e^{-y_2/3.0}, \qquad y_2 \ge 0,$$

an exponential density with mean 3.0. So $P(Y_2 > 6) = e^{-6/3.0} = 0.1353$.

(c) Definition 5.7, for $y_1 > 0$:

$$f(y_2 \mid y_1) = \frac{e^{-y_1/3.0}/9.00}{y_1 e^{-y_1/3.0}/9.00} = \frac{1}{y_1}, \qquad 0 \le y_2 \le y_1.$$

Given the total time, the queue time is uniform on $(0, y_1)$, so $P(Y_2 < 3.5 \mid Y_1 = 8) = 3.5/8 = 0.4375$.

(d) Definition 5.7, for $y_2 \ge 0$:

$$f(y_1 \mid y_2) = \frac{e^{-y_1/3.0}/9.00}{e^{-y_2/3.0}/3.0} = \frac{1}{3.0} e^{-(y_1 - y_2)/3.0}, \qquad y_1 \ge y_2.$$

Given the queue time, the remaining time $Y_1 - Y_2$ is exponential with mean 3.0, whatever y_2 was:

$$P(Y_1 > 10 \mid Y_2 = 4) = \int_{10}^{\infty} \frac{1}{3.0} e^{-(y_1-4)/3.0}\, dy_1 = e^{-6/3.0} = 0.1353.$$

Problem 14

seed 20260925

Each morning a bank's treasury allocates a trading desk a fraction Y_2 of the desk's maximum position limit; over many days Y_2 is uniformly distributed on $(0,1)$. Given the allocation $Y_2 = y_2$, the fraction of the maximum limit the desk actually uses that day, Y_1, is uniformly distributed on $(0, y_2)$.

Give every answer to at least four significant figures.

(a) The joint density of (Y_1, Y_2) at the point $y_1 = 0.25$, $y_2 = 0.40$: $f(0.25, 0.40) =$ ____________

(b) $P(Y_1 > 0.32 \mid Y_2 = 0.60) =$ ____________

(c) The unconditional probability $P(Y_1 \le 0.35) =$ ____________

(d) On a day when the desk used 0.40 of the maximum limit, the probability that it had been allocated more than 0.85: $P(Y_2 > 0.85 \mid Y_1 = 0.40) =$

____

(a) Definition 5.7 read backwards: $f(y_1, y_2) = f(y_1 \mid y_2)\, f_2(y_2)$. Here $f_2(y_2) = 1$ on $(0,1)$ and $f(y_1 \mid y_2) = 1/y_2$ on $(0, y_2)$, so

$$f(y_1, y_2) = \begin{cases} 1/y_2, & 0 < y_1 < y_2 < 1, \\ 0, & \text{elsewhere.} \end{cases}$$

At the given point $0.25 < 0.40$, so $f = 1/0.40 = 2.5000$.

(b) Straight from the conditional uniform density: $P(Y_1 > 0.32 \mid Y_2 = 0.60) = (0.60 - 0.32)/0.60 = 0.4667$.

(c) Definition 5.4(b): for $0 < y_1 < 1$, y_2 runs from y_1 to 1,

$$f_1(y_1) = \int_{y_1}^{1} \frac{1}{y_2}\, dy_2 = -\ln y_1.$$

Then, integrating by parts ($\int \ln y\, dy = y \ln y - y$),

$$P(Y_1 \le 0.35) = \int_0^{0.35} (-\ln y_1)\, dy_1 = 0.35\,(1 - \ln 0.35) = 0.35\,(1 - (-1.0498)) = 0.7174.$$

(d) Definition 5.7 in the other direction: for $0 < y_1 < 1$,

$$f(y_2 \mid y_1) = \frac{f(y_1, y_2)}{f_1(y_1)} = \frac{1/y_2}{-\ln y_1}, \qquad y_1 < y_2 < 1.$$

So

$$P(Y_2 > 0.85 \mid Y_1 = 0.40) = \frac{1}{-\ln 0.40} \int_{0.85}^{1} \frac{dy_2}{y_2} = \frac{-\ln 0.85}{-\ln 0.40} = \frac{0.1625}{0.9163} = 0.1774.$$

The model was stated “top down” (allocation, then usage given allocation); the marginal and the reverse conditional had to be derived. Observing heavy usage is evidence of a large allocation, and (d) quantifies how strong.

Problem 15

seed 20260925

A corporate treasury hedges its foreign-currency receivables. For a randomly chosen month let Y_1 be the gross open exposure and Y_2 the net exposure after hedging (both in million USD). Gross exposure never exceeds 19, and policy requires the hedge to cover enough that $2.0\,Y_2 \le Y_1$. The risk team models (Y_1, Y_2) as uniformly distributed over the region the policy allows:

$$f(y_1, y_2) = \begin{cases} \dfrac{2(2.0)}{19^2}, & 0 \le y_1 \le 19,\ 0 \le y_2,\ 2.0\,y_2 \le y_1, \\ 0, & \text{elsewhere.} \end{cases}$$

(a) For $0 \le y_2 \le 19/2.0$, the marginal density of net exposure $f_2(y_2) =$ ______________ (a formula in y_2)

(b) $P(Y_2 > 0.60) =$ ______________

(c) In a month when net exposure was 7.60, the probability that gross exposure exceeded 18.35: $P(Y_1 > 18.35 \mid Y_2 = 7.60) =$ ______________

Give probabilities to at least four significant figures.

The region is the triangle with vertices $(0, 0)$, $(19, 0)$ and $(19, 19/2.0)$; its area is $\frac{1}{2} \cdot 19 \cdot 19/2.0 = 361/(2 \cdot 2.0)$, which is why the constant height is $2(2.0)/19^2 = 0.01108$.

(a) Definition 5.4(b). Fix y_2 with $0 \le y_2 \le 19/2.0 = 9.5000$. The density is non-zero for $2.0\,y_2 \le y_1 \le 19$, so the lower limit depends on y_2:

$$f_2(y_2) = \int_{2.0 y_2}^{19} \frac{2(2.0)}{19^2}\, dy_1 = \frac{2(2.0)(19 - 2.0\,y_2)}{19^2}.$$

(b)

$$P(Y_2 > 0.60) = \int_{0.60}^{19/2.0} \frac{2(2.0)(19 - 2.0\,y_2)}{19^2}\, dy_2 = \left[-\frac{(19 - 2.0\,y_2)^2}{19^2}\right]_{0.60}^{19/2.0} = \left(\frac{19 - 1.200}{19}\right)^2 = 0.8777.$$

(c) Definition 5.7: for $0 \le y_2 < 19/2.0$,

$$f(y_1 \mid y_2) = \frac{2(2.0)/19^2}{2(2.0)(19 - 2.0\,y_2)/19^2} = \frac{1}{19 - 2.0\,y_2}, \qquad 2.0\,y_2 \le y_1 \le 19.$$

Given net exposure 7.60, gross exposure is uniform on $(15.200, 19)$, so

$$P(Y_1 > 18.35 \mid Y_2 = 7.60) = \frac{19 - 18.35}{19 - 15.200} = 0.1711.$$

Problem 16

seed 20260925

A regulator samples private pension funds at random. For a fund let Y_1 be the share of assets held in equities and Y_2 the share held in bonds; the rest is cash and property, so $Y_1 + Y_2 \le 1$. Across funds the joint density is

$$f(y_1,y_2)=\begin{cases}6\,y_1, & y_1\ge 0,\ y_2\ge 0,\ y_1+y_2\le 1,\\ 0, & \text{elsewhere.}\end{cases}$$

(a) For $0\le y_2\le 1$, $f_2(y_2)=$ ______________ (a formula in y_2)

(b) For a fund with exactly 0.15 in bonds, the probability of at least 0.40 in equities: $P(Y_1\ge 0.40\mid Y_2=0.15)=$ ______________

(c) For a fund with **at most** 0.20 in bonds, the probability of at least 0.50 in equities: $P(Y_1\ge 0.50\mid Y_2\le 0.20)=$ ______________

Give probabilities to at least four significant figures.

(a) Definition 5.4(b). For fixed $y_2\in[0,1]$, y_1 runs from 0 to $1-y_2$:

$$f_2(y_2)=\int_0^{1-y_2}6\,y_1\,dy_1=3(1-y_2)^2,\qquad 0\le y_2\le 1.$$

(b) Definition 5.7: for $0\le y_2<1$,

$$f(y_1\mid y_2)=\frac{6\,y_1}{3(1-y_2)^2}=\frac{2\,y_1}{(1-y_2)^2},\qquad 0\le y_1\le 1-y_2.$$

The support of the conditional density shrinks as y_2 grows. With $y_2=0.15$ the upper end is 0.85:

$$P(Y_1\ge 0.40\mid Y_2=0.15)=\int_{0.40}^{0.85}\frac{2\,y_1}{0.85^2}\,dy_1=1-\frac{0.40^2}{0.85^2}=0.7785.$$

(c) An event of positive probability, so $P(A\mid B)=P(A\cap B)/P(B)$:

$$P(Y_2\le 0.20)=\int_0^{0.20}3(1-y_2)^2\,dy_2=1-0.80^3=0.48800.$$

Because $0.50\le 1-0.20$, for every $y_2\le 0.20$ the equity share can range from 0.50 up to $1-y_2$:

$$P(Y_1\ge 0.50,Y_2\le 0.20)=\int_0^{0.20}\int_{0.50}^{1-y_2}6\,y_1\,dy_1\,dy_2=\int_0^{0.20}3\left[(1-y_2)^2-0.50^2\right]dy_2=0.48800-3(0.50)^2(0.20)=0.48800-0.15000=0.33800.$$

So $P(Y_1\ge 0.50\mid Y_2\le 0.20)=0.33800/0.48800=0.6926$.

Part (b) conditions on one value of Y_2 and uses Definition 5.7; part (c) averages over a whole range of bond shares and uses the elementary definition. They answer different questions and generally give different numbers.

Problem 17

seed 20260925

Electricity distributors in a region set an off-peak tariff Y_1 and a peak tariff Y_2 (both in AZN per 10 kWh). The peak tariff is never below the off-peak one, and the regulator caps the sum of the two at 2. Across distributors the tariff pair has joint density

$$f(y_1,y_2)=\begin{cases}6\,y_1^2\,y_2, & 0\le y_1\le y_2,\ y_1+y_2\le 2,\\ 0, & \text{elsewhere.}\end{cases}$$

Give probabilities to at least four significant figures.

(a) For $0\le y_1\le 1$, $f_1(y_1)=$ ______________ (a formula in y_1)

(b) $P(Y_2\le 0.75)=$ ______________

(c) $P(Y_2<0.93\mid Y_1=0.65)=$ ______________

(d) $P(Y_1\le 0.23\mid Y_2=1.45)=$ ______________

Sketch the support first: it is the triangle with vertices $(0,0)$, $(0,2)$ and $(1,1)$, bounded below by $y_2=y_1$ and above by $y_2=2-y_1$.

(a) Definition 5.4(b). For fixed $0\le y_1\le 1$, y_2 runs from y_1 to $2-y_1$:

$$f_1(y_1)=\int_{y_1}^{2-y_1}6y_1^2y_2\,dy_2=3y_1^2\left[(2-y_1)^2-y_1^2\right]=12\,y_1^2(1-y_1),$$

a beta density with $\alpha=3$, $\beta=2$.

(b) For fixed y_2, y_1 runs from 0 to whichever of y_2 and $2-y_2$ is smaller, so the marginal of Y_2 is piecewise:

$$f_2(y_2)=\begin{cases}\int_0^{y_2}6y_1^2y_2\,dy_1=2y_2^4, & 0\le y_2\le 1,\\ \int_0^{2-y_2}6y_1^2y_2\,dy_1=2y_2(2-y_2)^3, & 1<y_2\le 2.\end{cases}$$

$\int_0^1 2y_2^4\,dy_2=0.4$, and substituting $v=2-y_2$, $\int_1^u 2y_2(2-y_2)^3\,dy_2=\left[v^4-0.4v^5\right]_{2-u}^{1}=0.6-(2-u)^4+0.4(2-u)^5$.

$$\text{Since } 0.75\le 1:\quad P(Y_2\le 0.75)=\int_0^{0.75}2y_2^4\,dy_2=0.4(0.75)^5=0.0949.$$

(c) Definition 5.7: for $0 < y_1 < 1$,

$$f(y_2 \mid y_1) = \frac{6y_1^2 y_2}{12y_1^2(1-y_1)} = \frac{y_2}{2(1-y_1)}, \qquad y_1 \le y_2 \le 2 - y_1.$$

With $y_1 = 0.65$ the support is $[0.65, 1.35]$, and 0.93 lies inside it:

$$P(Y_2 < 0.93 \mid Y_1 = 0.65) = \int_{0.65}^{0.93} \frac{y_2}{2(0.35)}\, dy_2 = \frac{0.93^2 - 0.65^2}{4(0.35)} = 0.3160.$$

(d) Definition 5.7 with the piecewise f_2: Here the observed peak tariff exceeds 1, so the second piece applies and the conditional support is from 0 to 2 - y2: the regulator's cap, not the ordering, is what binds.

$$f(y_1 \mid y_2) = \frac{6y_1^2 y_2}{2y_2(2-y_2)^3} = \frac{3y_1^2}{(2-y_2)^3}, \quad 0 \le y_1 \le 2 - y_2$$

so, with $\min(y_2, 2-y_2) = 0.55$,

$$P(Y_1 \le 0.23 \mid Y_2 = 1.45) = \int_0^{0.23} \frac{3y_1^2}{0.55^3}\, dy_1 = \left(\frac{0.23}{0.55}\right)^3 = 0.0731.$$

Problem 18

seed 20260925

A property insurer writes policies whose limits (maximum payouts) range from 15 to 100 thousand AZN. For a randomly chosen claim, the limit Y_2 of the policy it came from is uniformly distributed on $(15, 100)$, and given $Y_2 = y_2$ the claim amount Y_1 (thousand AZN) is uniformly distributed on $(0, y_2)$.

Give every answer to at least four significant figures.

(a) For a policy with limit 49.0, the probability that the claim exceeds 13.5: $P(Y_1 > 13.5 \mid Y_2 = 49.0) =$ ____________

(b) The marginal density of the claim amount at $y_1 = 7.5$: $f_1(7.5) =$ ____________

(c) $P(Y_1 > 83.0) =$ ____________

(d) A claim of 7.5 thousand AZN has arrived. The probability that the policy's limit exceeds 74.5: $P(Y_2 > 74.5 \mid Y_1 = 7.5) =$ ____________

(a) Given the limit, the claim is uniform on $(0, 49.0)$, so $P(Y_1 > 13.5 \mid Y_2 = 49.0) = 1 - 13.5/49.0 = 0.7245$.

Joint density. Definition 5.7 rearranged, $f(y_1, y_2) = f(y_1 \mid y_2)\, f_2(y_2) = \frac{1}{y_2} \cdot \frac{1}{85}$, on the region $15 < y_2 < 100$, $0 < y_1 < y_2$.

(b) Definition 5.4(b): for fixed y_1, y_2 must satisfy **both** $y_2 > 15$ and $y_2 > y_1$, so the lower limit is $\max(15, y_1)$:

$$f_1(y_1) = \int_{\max(15,y_1)}^{100} \frac{dy_2}{85\, y_2} = \begin{cases} \dfrac{\ln(100/15)}{85}, & 0 < y_1 \le 15, \\ \dfrac{\ln(100/y_1)}{85}, & 15 < y_1 < 100. \end{cases}$$

The claim density is flat up to the smallest limit and then decreases. Since $7.5 < 15$, $f_1(7.5) = 1.89712/85 = 0.022319$.

(c) 83.0 lies on the second piece. Using $\int \ln y\, dy = y \ln y - y$,

$$P(Y_1 > 83.0) = \frac{1}{85} \int_{83.0}^{100} (\ln 100 - \ln y_1)\, dy_1 = \frac{(100 - 83.0) - 83.0 \ln(100/83.0)}{85} = \frac{17.0 - 83.0(0.18633)}{85} = 0.0181.$$

(d) Definition 5.7 the other way round: for $0 < y_1 \le 15$,

$$f(y_2 \mid y_1) = \frac{f(y_1, y_2)}{f_1(y_1)} = \frac{1/(85\, y_2)}{\ln(100/15)/85} = \frac{1}{y_2 \ln(100/15)}, \qquad 15 < y_2 < 100.$$

Hence

$$P(Y_2 > 74.5 \mid Y_1 = 7.5) = \frac{1}{\ln(100/15)} \int_{74.5}^{100} \frac{dy_2}{y_2} = \frac{\ln(100/74.5)}{\ln(100/15)} = \frac{0.29437}{1.89712} = 0.1552.$$

Notice that this does not depend on the claim amount as long as it is below 15: every limit in the range could have produced such a small claim, and smaller limits make any particular small claim more likely (density $1/y_2$). The prior share of limits above 74.5 would be $(100 - 74.5)/85 = 0.3000$, larger than 0.1552: the claim shifts belief towards smaller limits.

Problem 19

seed 20260925

A bank's anti-fraud system raises alerts at two branches. The weekly numbers of alerts, Y_1 at branch 1 and Y_2 at branch 2, are independent Poisson random variables with means $\lambda_1 = 2.0$ and $\lambda_2 = 6.0$. Let $W = Y_1 + Y_2$; you may use the fact (proved in Chapter 6) that W is Poisson with mean $\lambda_1 + \lambda_2$.

This week there were $W = 6$ alerts in total. Give every probability to at least four significant figures.

(a) Given $W = 6$, the conditional distribution of Y_1 is binomial with $n = 6$ and success probability $p =$ ______________

(b) $P(Y_1 = 3 \mid W = 6) =$ ______________

(c) The probability that branch 1 raised at least half of this week's alerts, $P(Y_1 \geq 3 \mid W = 6) =$ ______________

(a) Definition 5.5 with the conditioning variable W. For $y_1 = 0, 1, \ldots, w$, the event $\{Y_1 = y_1, W = w\}$ is the event $\{Y_1 = y_1, Y_2 = w - y_1\}$, so by independence

$$P(Y_1 = y_1 \mid W = w) = \frac{P(Y_1 = y_1)P(Y_2 = w - y_1)}{P(W = w)} = \frac{\dfrac{\lambda_1^{y_1} e^{-\lambda_1}}{y_1!} \cdot \dfrac{\lambda_2^{w-y_1} e^{-\lambda_2}}{(w-y_1)!}}{\dfrac{(\lambda_1+\lambda_2)^w e^{-(\lambda_1+\lambda_2)}}{w!}} = \binom{w}{y_1} \left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^{y_1} \left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{w-y_1}.$$

The exponentials cancel, and what is left is the binomial probability function with $n = w$ and $p = \lambda_1/(\lambda_1 + \lambda_2) = 2.0/8.0 = 0.25000$.

(b) $P(Y_1 = 3 \mid W = 6) = \binom{6}{3}(0.25000)^3(0.75000)^3 = 20 \times (0.25000)^3(0.75000)^3 = 0.1318$.

(c) At least half of 6 alerts means $Y_1 \geq 3$. Summing the conditional (binomial) probabilities $\binom{6}{y_1}(0.25000)^{y_1}(0.75000)^{6-y_1}$ for $y_1 = 3, \ldots, 6$:

$$P(Y_1 \geq 3 \mid W = 6) = 0.1318 + 0.0330 + 0.0044 + 0.0002 = 0.1694.$$

Compare the unconditional $P(Y_1 = 3) = e^{-2.0}(2.0)^3/3! = 0.1804$. Once the week's total is known, the only uncertainty left is how the alerts split between branches, and each alert independently lands at branch 1 with probability equal to that branch's share of the total rate. An investigator who sees an unusually high total therefore should not conclude that **both** branches had a bad week; the conditional split is what tells her where to look.

Problem 20

seed 20260925

A credit analyst is unsure of the one-year default probability Θ of small firms in the retail sector. She models it as a random variable with density

$$f(\theta) = 9(1-\theta)^8, \qquad 0 < \theta < 1,$$

(a beta density with $\alpha = 1$, $\beta = 9$, mean 0.1000). A bank holds loans to 28 such firms; given $\Theta = \theta$, the number Y of them that default is binomial with $n = 28$ and probability θ, so the joint distribution of (Y, Θ) is $p(y \mid \theta)\, f(\theta)$.

Give every probability to at least four significant figures.

(a) Before any data, $P(\Theta > 0.11) =$ ______________

(b) The marginal probability that none of the 28 firms defaults, $P(Y = 0) =$ ______________

(c) The marginal probability of exactly one default, $P(Y = 1) =$ ______________

(d) A year passes with no defaults. The conditional probability that the sector's default probability exceeds 0.11: $P(\Theta > 0.11 \mid Y = 0) =$ ______________

(a) $P(\Theta > 0.11) = \int_{0.11}^1 9(1-\theta)^8\, d\theta = (1 - 0.11)^9 = 0.89^9 = 0.3504$.

(b) The marginal probability function of Y integrates θ out of the joint distribution (the mixed discrete-continuous analogue of Definition 5.4):

$$P(Y = 0) = \int_0^1 (1-\theta)^{28} \cdot 9(1-\theta)^8\, d\theta = 9 \int_0^1 (1-\theta)^{36}\, d\theta = \frac{9}{37} = 0.2432.$$

(c) Similarly, using $\int_0^1 \theta(1-\theta)^m\, d\theta = \frac{1}{(m+1)(m+2)}$ (a beta integral, or integration by parts) with $m = 37 - 2$:

$$P(Y = 1) = \int_0^1 28\,\theta(1-\theta)^{28-1} \cdot 9(1-\theta)^8\, d\theta = 28 \cdot 9 \int_0^1 \theta(1-\theta)^{37-2}\, d\theta = \frac{28 \cdot 9}{36 \cdot 37} = 0.1892.$$

(d) The conditional density of Θ given $Y = 0$ is the joint divided by the marginal of the conditioning variable, exactly as in Definition 5.7:

$$f(\theta \mid Y = 0) = \frac{(1-\theta)^{28} \cdot 9(1-\theta)^8}{9/37} = 37(1-\theta)^{36}, \qquad 0 < \theta < 1,$$

a beta density with $\alpha = 1$, $\beta = 37$. So

$$P(\Theta > 0.11 \mid Y = 0) = (1 - 0.11)^{37} = 0.0134,$$

down from 0.3504 before the data. A clean year is evidence that the sector is safer than feared, and the model says precisely how much: each default-free loan multiplies the tail probability by another factor 0.89.

Answer key

Problem	Answers

1	(a) 0.15, (b) 0.26, (c) 0.16, (d) 0.27, (e) 0.76
2	(a) 0.09, (b) 0.105263, (c) 0.222222, (d) 0.95
3	(a) 0.6*y_1+0.7, (b) 0.81175, (c) 0.768
4	(a) 0.729167, (b) 0.104167, (c) 0.104167, (d) 0.461538
5	(a) 0.000408163, (b) 0.0228571, (c) 0.232143, (d) 0.185714
6	(a) 0.333333*e^(-y_1/3), (b) 0.398858, (c) 0.398858, (d) 0.27414
7	(a) 0.12, (b) 0.32, (c) 0.56, (d) 0.518519, (e) 0.176471
8	(a) 0.149035, (b) 0.0607211, (c) 0.516129, (d) 0.165217
9	(a) 0.0245098, (b) 0.533333, (c) 0.5, (d) 0.147059
10	(a) 1.5*(1-y_2^2), (b) 0.722222, (c) 0.639498
11	(a) 3*(1-y_1)^2, (b) 0.617347, (c) 0.378698
12	(a) 0.777778, (b) 0.64, (c) 0.805
13	(a) y_1*e^(-y_1/3)/9, (b) 0.135335, (c) 0.4375, (d) 0.135335
14	(a) 2.5, (b) 0.466667, (c) 0.717438, (d) 0.177366
15	(a) 4*(19-2*y_2)/361, (b) 0.877673, (c) 0.171053
16	(a) 3*(1-y_2)^2, (b) 0.778547, (c) 0.692623
17	(a) 12*y_1^2*(1-y_1), (b) 0.0949219, (c) 0.316, (d) 0.07313
18	(a) 0.72449, (b) 0.0223191, (c) 0.0180546, (d) 0.155167
19	(a) 0.25, (b) 0.131836, (c) 0.169434
20	(a) 0.350356, (b) 0.243243, (c) 0.189189, (d) 0.01341