Week 5, Problem Set 2 - worked solutions

Wackerly 3.1-3.3 - discrete random variables and expected value

> 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 microfinance lender has four small loans outstanding to market traders. Let Y be the number of these loans that default within the year. The lender's risk team models Y with the probability distribution below, but the entry for $y = 0$ has been left blank as an unknown constant c.

y	0	1	2	3	4
$p(y)$	c	0.20	0.14	0.03	0.03

Give every probability to at least four significant figures.

(a) The value of c that makes this a valid probability distribution: $c =$ ____________

(b) The probability that at least two loans default: $P(Y \geq 2) =$ ____________

(c) $P(1 \leq Y \leq 3) =$ ____________

(a) Theorem 3.1 says that every $p(y)$ lies in $[0, 1]$ and that the probabilities over all values with nonzero probability sum to 1:

$$c + 0.20 + 0.14 + 0.03 + 0.03 = 1 \quad \Longrightarrow \quad c = 1 - 0.40 = 0.60.$$

The value lies in $[0, 1]$, so the first condition of Theorem 3.1 also holds and the table is a genuine probability distribution. Note that c is the probability that **no** loan defaults, the number the lender cares about most.

(b) The event $(Y \geq 2)$ is the union of the disjoint events $(Y = 2)$, $(Y = 3)$, $(Y = 4)$, so its probability is the sum of their probabilities:

$$P(Y \geq 2) = p(2) + p(3) + p(4) = 0.14 + 0.03 + 0.03 = 0.20.$$

(c) Both endpoints are included, so three values count:

$$P(1 \leq Y \leq 3) = p(1) + p(2) + p(3) = 0.20 + 0.14 + 0.03 = 0.37.$$

As a check, $P(1 \leq Y \leq 3) = 1 - p(0) - p(4) = 1 - 0.60 - 0.03$, which gives the same number.

Problem 2

seed 20260925

A mobile operator's call centre in Baku logs complaints about dropped calls. Let Y be the number of such complaints received in a randomly chosen fifteen-minute interval. Its probability distribution is

y	0	1	2	3	4	5
$p(y)$	0.05	0.18	0.41	0.17	0.14	0.05

and $p(y) = 0$ for every other y.

Give every probability to at least four significant figures.

(a) $P(1 \le Y \le 4) =$ ______________

(b) $P(Y > 4) =$ ______________

(c) $P(1 < Y < 4) =$ ______________

Every event here is a union of the disjoint events $(Y = y)$, so by Definition 3.2 its probability is the sum of $p(y)$ over the values of y it contains. The only work is reading the inequality signs.

(a) With $\le$ on both sides, both endpoints are in: $y = 1, 2, 3, 4$, and

$$P(1 \le Y \le 4) = \sum_{y=1}^{4} p(y) = 0.90.$$

(b) "More than 4" means $y = 5$: $P(Y > 4) = 0.05$.

(c) With strict inequalities both endpoints drop out, leaving $y = 2, 3$:

$$P(1 < Y < 4) = \sum_{y=2}^{3} p(y) = 0.58.$$

For a continuous quantity the difference between (a) and (c) would not matter. For a discrete random variable each endpoint carries its own positive probability, here $p(1)$ and $p(4)$, and the two answers differ by exactly those two table entries.

Problem 3

seed 20260925

An exporter in Ganja sends a payment abroad through its bank. Let Y be the number of business days until the payment settles. The bank's operations team models Y with the probability function

$$p(y) = c\,(14 - y), \qquad y = 1, 2, \ldots, 4,$$

and $p(y) = 0$ otherwise, where c is a constant.

Give every answer to at least four significant figures (a fraction such as $1/37$ is also accepted in (a)).

(a) $c =$ ______________

(b) $P(Y \le 2) =$ ______________

(c) $E(Y) =$ ______________ business days

(d) While the payment is in transit the exporter loses 9 AZN per business day in interest. The expected interest cost of one payment is ______________ AZN.

(a) The weights $14 - y$ are positive for $y = 1, \ldots, 4$, so $p(y) \ge 0$ as long as $c > 0$. The second condition of Theorem 3.1 fixes c:

$$\sum_{y=1}^{4} c\,(14 - y) = c\,\big(13 + 12 + \cdots + 10\big) = c \cdot 46 = 1 \quad\Longrightarrow\quad c = \frac{1}{46} = 0.021739.$$

(b) $P(Y \le 2) = p(1) + p(2) = c(13 + 12) = 25/46 = 0.5435.$

(c) By Definition 3.4,

$$E(Y) = \sum_{y=1}^{4} y\,p(y) = c\sum_{y=1}^{4} y\,(14-y) = \frac{110}{46} = 2.3913.$$

(d) The cost is $g(Y) = 9Y$. By Theorem 3.4, $E(9Y) = 9\,E(Y) = 9 \times 2.3913 = 21.52$ AZN.

There is no need to build the distribution of the cost first: a constant multiple comes straight out of the expectation.

Problem 4

seed 20260925

An insurer in Baku sells compulsory motor third-party liability cover. For a randomly chosen policy let Y be the number of claims filed during the policy year. From past portfolio data,

y	0	1	2	3
$p(y)$	0.83	0.10	0.06	0.01

Each claim costs the insurer 2200 AZN on average, and for this problem treat every claim as costing exactly that amount.

(a) The expected number of claims per policy: $E(Y) =$ ______________

(b) The expected claim cost per policy: ______________ AZN

(c) The probability that a policy produces at least one claim: ______________

(a) By Definition 3.4,

$$E(Y) = \sum_{y} y\,p(y) = 0(0.83) + 1(0.10) + 2(0.06) + 3(0.01) = 0.25.$$

The $y = 0$ term contributes nothing, however large $p(0)$ is: most policies are claim-free, and the mean is pulled up only by the few that are not.

(b) The claim cost is $2200Y$, a constant times Y. By Theorem 3.4,

$$E(2200\,Y) = 2200\,E(Y) = 2200 \times 0.25 = 550.00 \text{ AZN.}$$

This is the **pure premium**: what the insurer must collect per policy, on average, just to pay claims, before any expenses or profit.

(c) $P(Y \geq 1) = 1 - p(0) = 1 - 0.83 = 0.17$.

Comparing (a) and (c): $E(Y) = 0.25$ is larger than $P(Y \geq 1) = 0.17$ because a policy with two or three claims counts two or three times in the mean but only once in the probability.

Problem 5

seed 20260925

An agricultural insurer offers vineyard owners near Shamakhi a one-year hail policy that pays 5500 AZN if hail destroys the harvest and nothing otherwise. Historically a hail loss of this kind occurs with probability 0.047 per vineyard per year. Handling each policy costs the insurer 20 AZN in administration, whether or not a claim is paid. The insurer's profit on a policy is what remains of the premium after payouts and administration.

(a) The expected payout on one policy: ______________ AZN

(b) The premium C the insurer must charge so that its expected profit per policy is 75 AZN: $C =$ ______________ AZN

(c) The sales team proposes a premium of 285 AZN instead. The expected profit per policy at that premium is ____________ AZN (enter a negative number for an expected loss).

(a) The payout X takes the value 5500 with probability 0.047 and 0 with probability 0.953. By Definition 3.4,

$$E(X) = 5500(0.047) + 0(0.953) = 258.50 \text{ AZN}.$$

(b) If the premium is C, the profit is

$$\text{profit} = \begin{cases} C - 20 & \text{if no hail (probability 0.953)}, \\ C - 20 - 5500 & \text{if hail (probability 0.047)}, \end{cases}$$

that is, profit $= C - 20 - X$. By Theorems 3.3 and 3.5, $E(\text{profit}) = C - 20 - E(X) = C - 20 - 258.50$. Setting this equal to 75:

$$C = 75 + 20 + 258.50 = 353.50 \text{ AZN}.$$

(c) With $C = 285$, $E(\text{profit}) = 285 - 20 - 258.50 = 6.50$ AZN.

Notice what the expected profit does **not** say: in a hail year the insurer earns $265.00 - 5500 = -5235.00$ AZN on this policy. The expectation is a long-run average over many vineyards, which is why the premium only works for an insurer that holds a large portfolio.

Problem 6

seed 20260925

A household in Sheki receives a monthly remittance of Y US dollars from a relative working abroad, where $E(Y) = 420$ and the standard deviation of Y is 285 USD. The receiving bank converts the transfer at a fixed rate of 1.69 AZN per USD and deducts a flat fee of 9 AZN, so the household is credited

$$X = 1.69\,Y - 9 \quad \text{AZN}.$$

Give every answer to at least four significant figures.

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

(b) $V(X) =$ ____________ AZN^2

(c) The standard deviation of X: ____________ AZN

Write $X = aY + b$ with $a = 1.69$ and $b = -9$.

(a) By Theorem 3.5 the expectation of a sum is the sum of the expectations; by Theorem 3.4 $E(aY) = aE(Y)$; by Theorem 3.3 $E(b) = b$. So

$$E(X) = aE(Y) + b = 1.69(420) - 9 = 709.80 - 9 = 700.80.$$

(b) By Definition 3.5, $V(X) = E\{[X - E(X)]^2\}$. Here $X - E(X) = (aY + b) - (a\mu + b) = a(Y - \mu)$: the fee cancels. Hence, by Theorem 3.4,

$$V(X) = E[a^2(Y-\mu)^2] = a^2E[(Y-\mu)^2] = a^2V(Y) = 2.8561 \times 81225 = 231986.72.$$

(c) $\sigma_X = \sqrt{V(X)} = |a|\,\sigma_Y = 1.69 \times 285 = 481.65$ AZN.

The flat fee moves every outcome by the same amount, so it shifts the mean and leaves the spread untouched. The exchange rate stretches every deviation from the mean by the factor 1.69, so the standard deviation is multiplied by 1.69 and the variance by its square.

Problem 7

seed 20260925

A pension fund holds four corporate bonds. Let Y be the number of them whose credit rating is downgraded during the next year. The fund's credit analysts use the distribution

y	0	1	2	3	4
$p(y)$	0.52	0.22	0.18	0.05	0.03

Give every answer to at least four significant figures.

(a) $\mu = E(Y) =$ ______________

(b) $E(Y^2) =$ ______________

(c) $\sigma^2 = V(Y) =$ ______________

(d) $\sigma =$ ______________

(a) Definition 3.4:

$$\mu = \sum_{y=0}^{4} y\,p(y) = 0(0.52) + 1(0.22) + 2(0.18) + 3(0.05) + 4(0.03) = 0.85.$$

(b) Theorem 3.2 with $g(Y) = Y^2$:

$$E(Y^2) = \sum_{y=0}^{4} y^2 p(y) = 0(0.52) + 1(0.22) + 4(0.18) + 9(0.05) + 16(0.03) = 1.87.$$

(c) Theorem 3.6:

$$\sigma^2 = E(Y^2) - \mu^2 = 1.87 - 0.85^2 = 1.87 - 0.7225 = 1.1475.$$

The same number comes out of Definition 3.5 directly, one term per value of y:

$$E[(Y-\mu)^2] \approx 0.3757 + 0.0050 + 0.2380 + 0.2311 + 0.2977 \approx 1.1475.$$

Theorem 3.6 is the shorter route because y^2 is an integer for every y, while $(y-\mu)^2$ is not.

(d) The standard deviation is the positive square root, $\sigma = \sqrt{1.1475} = 1.0712$ downgrades. It is measured in the same units as Y, which is why it, and not the variance, is the number to quote next to the mean.

Problem 8

seed 20260925

A venture-debt fund lends to a Baku software start-up. The loan is repaid in a single payment of 190 thousand AZN at the end of year Y, where the fund models the repayment year as

y	1	2	3	4
$p(y)$	0.18	0.35	0.19	0.28

The fund discounts at 10% per year, so the present value of the repayment is

$$g(Y) = \frac{190}{(1.10)^Y} \quad \text{thousand AZN.}$$

Give every answer to at least four significant figures.

(a) $E(Y) =$ ______________ years

(b) The expected present value $E[g(Y)] =$ ______________ thousand AZN

(c) The present value computed at the mean repayment year, $g(E(Y)) = 190/(1.10)^{E(Y)} =$ ______________ thousand AZN

(a) Definition 3.4: $E(Y) = 1(0.18) + 2(0.35) + 3(0.19) + 4(0.28) = 2.57$.

(b) By Theorem 3.2, $E[g(Y)] = \sum_y g(y)\,p(y)$, so first evaluate g at each possible year:

$$g(1) = 172.7273, \quad g(2) = 157.0248, \quad g(3) = 142.7498, \quad g(4) = 129.7726.$$

Then weight each by its probability:

$$E[g(Y)] = 172.7273(0.18) + 157.0248(0.35) + 142.7498(0.19) + 129.7726(0.28) = 149.5084.$$

(c) $g(E(Y)) = 190/(1.10)^{2.57} = 148.7217$.

The two numbers differ: $E[g(Y)] - g(E(Y)) = 0.7866$ thousand AZN. Theorem 3.2 averages $g(y)$, not y, and for a function that is not linear the average of $g(y)$ is not g of the average. Here g is convex (it falls steeply at first and then flattens), so an early repayment raises present value by more than an equally late one lowers it, and $E[g(Y)] > g(E(Y))$ at every draw. A fund that plugs the expected repayment date into its discounting formula undervalues uncertain-timing loans.

Problem 9

seed 20260925

A bank's credit committee has 11 loan applications on the table this week: 2 from small and medium-sized enterprises (SMEs) and 9 from large corporates. It has time to review only 3, and to avoid any appearance of favouritism it draws the three at random, every set of three being equally likely. Let Y be the number of SME applications among the three drawn.

Give every answer to at least four significant figures (fractions are accepted).

(a) $p(0) = P(Y = 0) =$ ______________

(b) $p(1) =$ ______________

(c) $E(Y) =$ ______________

(d) A state programme pays the bank a subsidy of 450 AZN for every SME application that is reviewed. The expected subsidy from this week's draw is ______________ AZN.

The sample space is the set of all ways to choose 3 applications from 11: $\binom{11}{3} = 165$ equally likely sample points, each with probability $1/165$. By Definition 3.2, $p(y)$ is the number of sample points assigned the value y, divided by 165. To get $Y = y$ the committee must pick y of the 2 SME files and $3 - y$ of the 9 corporate ones, so

$$p(y) = \frac{\binom{2}{y}\binom{9}{3-y}}{\binom{11}{3}}, \qquad y = 0, 1, 2, 3.$$

(a) $p(0) = \binom{9}{3}/165 = 84/165 = 0.5091$.

(b) $p(1) = 2\binom{9}{2}/165 = 2 \cdot 36/165 = 0.4364$.

For the mean we also need $p(2) = \binom{2}{2}(9)/165 = 9/165 = 0.0545$ and $p(3) = \binom{2}{3}/165 = 0/165 = 0.0000$. The four probabilities sum to $(84 + 72 + 9 + 0)/165 = 1$, as Theorem 3.1 requires.

(c) Definition 3.4:

$$E(Y) = \frac{0(84) + 1(72) + 2(9) + 3(0)}{165} = \frac{90}{165} = 0.5455.$$

This equals $3 \times 2/11$: on average the three files drawn contain SMEs in the same proportion as the whole pile. (Section 3.7 will show why.)

(d) The subsidy is $450Y$, so by Theorem 3.4 its expectation is $450\,E(Y) = 450 \times 0.5455 = 245.45$ AZN.

Problem 10

seed 20260925

A carpet exporter in Guba sends cross-border wire transfers to suppliers. Let Y be the number of transfers it sends in a randomly chosen week:

y	1	2	3	4	5
$p(y)$	0.12	0.30	0.34	0.14	0.10

Its bank charges a fixed account fee of 7 AZN per week plus 5.50 AZN per transfer, so the weekly bank charge is $C = 7 + 5.50\,Y$ AZN.

Give every answer to at least four significant figures.

(a) $E(Y) =$ ______________

(b) $V(Y) =$ ______________

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

(d) The standard deviation of C: ______________ AZN

(a) Definition 3.4:

$$E(Y) = 1(0.12) + 2(0.30) + 3(0.34) + 4(0.14) + 5(0.10) = 2.80.$$

(b) Theorem 3.2 gives

$$E(Y^2) = 1(0.12) + 4(0.30) + 9(0.34) + 16(0.14) + 25(0.10) = 9.12,$$

and Theorem 3.6 gives $V(Y) = E(Y^2) - \mu^2 = 9.12 - 2.80^2 = 1.2800$.

(c) By Theorems 3.3, 3.4 and 3.5, $E(C) = E(7) + E(5.50Y) = 7 + 5.50\,E(Y) = 7 + 5.50(2.80) = 22.40$ AZN.

(d) $C - E(C) = 5.50\,(Y - \mu)$, so by Definition 3.5 and Theorem 3.4

$$V(C) = E\{5.50^2(Y-\mu)^2\} = 5.50^2\,V(Y) = 38.7200, \qquad \sigma_C = 5.50\,\sigma_Y = 5.50 \times 1.1314 = 6.2225 \text{ AZN}.$$

The fixed fee of 7 AZN is paid every week whatever happens, so it raises the expected charge but contributes nothing to its variability.

Problem 11

seed 20260925

An investment committee in Baku must fund one of two expansion projects. Their one-year returns (in percent) depend on the state of the regional economy:

State	Recession	Normal	Boom
Probability	0.18	0.53	0.29
Return of A, R_A (%)	-4	8	13
Return of B, R_B (%)	1	7	16

Give every answer to at least four significant figures.

(a) $E(R_A) =$ ______________ % $\quad E(R_B) =$ ______________ %

(b) Standard deviations: $\sigma_A =$ ______________ % $\quad \sigma_B =$ ______________ %

(c) An investor prefers a higher expected return and a lower standard deviation. Does one project dominate the other on both counts?

- ○ Project A: higher expected return and lower standard deviation
- ○ Project B: higher expected return and lower standard deviation
- ○ Neither: the project with the higher expected return is also the riskier one

Each return is a discrete random variable taking three values, one per state.

(a) Definition 3.4:

$$E(R_A) = (-4)(0.18) + (8)(0.53) + (13)(0.29) = 7.2900,$$
$$E(R_B) = (1)(0.18) + (7)(0.53) + (16)(0.29) = 8.5300.$$

(b) Theorem 3.2 with $g(R) = R^2$, then Theorem 3.6:

$$E(R_A^2) = (-4)^2(0.18) + (8)^2(0.53) + (13)^2(0.29) = 85.8100, \quad V(R_A) = 85.8100 - 7.2900^2 = 32.6659,$$
$$E(R_B^2) = (1)^2(0.18) + (7)^2(0.53) + (16)^2(0.29) = 100.3900, \quad V(R_B) = 100.3900 - 8.5300^2 = 27.6291.$$

So $\sigma_A = \sqrt{32.6659} = 5.7154$ and $\sigma_B = \sqrt{27.6291} = 5.2563$.

(c) Compare 7.2900 with 8.5300, and 5.7154 with 5.2563. The answer is: Project B: higher expected return and lower standard deviation.

When one project has both the higher mean and the lower standard deviation, every investor who likes return and dislikes risk picks it, and no attitude to risk needs to be specified. When the higher mean comes with the higher standard deviation, the mean and variance alone cannot settle the choice: the answer depends on how much expected return the investor demands for each extra point of risk.

Problem 12

seed 20260925

A freight forwarder tracks how long import containers wait for customs clearance at the Port of Baku in Alat. Let Y be the waiting time, in whole days, of a randomly chosen container:

y	1	2	3	4	5	6	7
$p(y)$	0.06	0.25	0.23	0.23	0.14	0.05	0.04

Give every answer to at least four significant figures.

(a) $\mu = E(Y) =$ _______________ days

(b) $\sigma =$ _______________ days

(c) The probability that Y falls strictly inside one standard deviation of the mean, $P(\mu - \sigma < Y < \mu + \sigma) =$

____

(d) $P(\mu - 2\sigma < Y < \mu + 2\sigma) =$ _______________

(a) Definition 3.4: $\mu = \sum_{y=1}^{7} y\,p(y) = 3.45$ days.

(b) Theorem 3.2 gives $E(Y^2) = \sum_{y=1}^{7} y^2 p(y) = 14.07$, so by Theorem 3.6

$$\sigma^2 = E(Y^2) - \mu^2 = 14.07 - 3.45^2 = 2.1675, \qquad \sigma = 1.4722.$$

(c) The interval is $(1.9778,\ 4.9222)$. Y takes only whole-day values, and the ones inside are $y = 2, 3, 4$. Adding their probabilities, $P(\mu - \sigma < Y < \mu + \sigma) = 0.71$.

(d) The interval is $(0.5055,\ 6.3945)$, which contains $y = 1, 2, 3, 4, 5, 6$, so $P(\mu - 2\sigma < Y < \mu + 2\sigma) = 0.96$.

How to read these numbers. The empirical rule of Chapter 1 suggests roughly 68% within one standard deviation and 95% within two for a mound-shaped distribution, and Example 3.2 shows it can work tolerably even for a small discrete one. Tchebysheff's theorem (Section 3.11) makes a promise that holds for **every** distribution: at least $1 - 1/2^2 = 0.75$ of the probability lies within two standard deviations, and the answer to (d) respects that floor. For the forwarder, σ is the natural yardstick for "an unusual delay": a wait longer than $\mu + 2\sigma \approx 6.3945$ days is, by part (d), an event of probability at most $1 - 0.96 = 0.04$.

Problem 13

seed 20260925

An insurer in Baku writes one-year fire policies on warehouses, each insured for 85000 AZN. From its loss history, a warehouse of this type is totally destroyed in a year with probability 0.0040 and suffers a partial loss of 60% of its value with probability 0.011. Ignore all other partial losses. Let X be the insurer's payout on one policy.

Give every answer to at least four significant figures.

(a) The premium at which the insurer breaks even on average, $E(X) =$ _______________ AZN

(b) Each policy also costs 40 AZN to administer, and the insurer wants an expected profit of 140 AZN per policy. The premium it should charge is _______________ AZN

(c) The standard deviation of the payout, $\sigma_X =$ _______________ AZN

The payout has three possible values:

x	85000	51000	0
$p(x)$	0.0040	0.011	0.9850

where $p(0) = 1 - 0.0040 - 0.011$ by Theorem 3.1.

(a) Definition 3.4:

$$E(X) = 85000(0.0040) + 51000(0.011) + 0 = 340.00 + 561.00 = 901.00 \text{ AZN.}$$

If the premium equals $E(X)$, the insurer's expected profit on a policy is zero: it breaks even in the long run over many warehouses.

(b) With premium C the profit is $C - 40 - X$. By Theorems 3.3 and 3.5 its expectation is $C - 40 - E(X)$, and setting this to 140 gives

$$C = E(X) + 40 + 140 = 901.00 + 40 + 140 = 1081.00 \text{ AZN.}$$

(c) Theorem 3.2 gives $E(X^2) = 85000^2(0.0040) + 51000^2(0.011) = 57511000.0$, and Theorem 3.6 gives

$$V(X) = E(X^2) - [E(X)]^2 = 57511000.0 - 901.00^2 = 56699199.0, \qquad \sigma_X = 7529.89 \text{ AZN.}$$

The standard deviation is many times the mean. A single policy is almost always a small gain and very occasionally a catastrophic loss, which is why a break-even premium only makes sense across a large book of policies, and why insurers hold capital against the spread and not just the mean.

Problem 14

seed 20260925

A data centre in Sumgait that hosts servers for several banks records Y, the number of unplanned power outages in a month:

y	0	1	2	3
$p(y)$	0.52	0.23	0.20	0.05

Because repeated outages trigger penalty clauses in its client contracts, the monthly cost of running the backup system is

$$C = 1400 + 400\,Y + 275\,Y^2 \quad \text{AZN.}$$

Give every answer to at least four significant figures.

(a) $E(Y) =$ ______________

(b) $E(Y^2) =$ ______________

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

(a) Definition 3.4: $E(Y) = 0(0.52) + 1(0.23) + 2(0.20) + 3(0.05) = 0.78$.

(b) Theorem 3.2 with $g(Y) = Y^2$: $E(Y^2) = 0(0.52) + 1(0.23) + 4(0.20) + 9(0.05) = 1.48$.

(c) C is a sum of three functions of Y. Theorem 3.5 splits the expectation, Theorem 3.3 handles the constant and Theorem 3.4 the constant multiples:

$$E(C) = 1400 + 400\,E(Y) + 275\,E(Y^2) = 1400 + 400(0.78) + 275(1.48) = 2119.00 \text{ AZN.}$$

Check through Theorem 3.2 directly: the cost is 1400, 2075, 3300, 5075 AZN at $y = 0, 1, 2, 3$, and $1400(0.52) + 2075(0.23) + 3300(0.20) + 5075(0.05) = 2119.00$.

A tempting shortcut is to plug the mean into the cost formula: $1400 + 400(0.78) + 275(0.78)^2 = 1879.31$. It is too low by $275\,[E(Y^2) - \mu^2] = 275\,V(Y) = 239.69$ AZN, because $E(Y^2) \neq [E(Y)]^2$. The quadratic penalty makes the variability of outages cost money in its own right, not just their average number.

Problem 15

seed 20260925

An analyst models next year's consumer-price inflation in Azerbaijan, Y (in percent), as a discrete random variable over a grid of scenarios, with $E(Y) = 4.4$ and standard deviation 2.8. The central bank's target is 3.5%, and a common way to score a policy outcome is the squared miss $(Y - 3.5)^2$.

Give every answer to at least four significant figures.

(a) $E(Y^2) =$ ______________

(b) The expected squared miss from the target, $E[(Y - 3.5)^2] =$ ______________

(c) Of all constants c, the one that makes $E[(Y - c)^2]$ as small as possible is $c =$ ______________, and the smallest possible value of $E[(Y - c)^2]$ is ______________

(a) Theorem 3.6 says $V(Y) = E(Y^2) - \mu^2$. Rearranged,

$$E(Y^2) = V(Y) + \mu^2 = 2.8^2 + 4.4^2 = 7.84 + 19.36 = 27.20.$$

(b) Expand the square and use Theorems 3.3, 3.4 and 3.5:

$$E[(Y - 3.5)^2] = E(Y^2) - 7.0\,E(Y) + 3.5^2 = 27.20 - 30.80 + 12.25 = 8.65.$$

(c) The same expansion for a general constant c gives

$$E[(Y - c)^2] = E(Y^2) - 2c\mu + c^2 = \big(E(Y^2) - \mu^2\big) + (\mu - c)^2 = \sigma^2 + (\mu - c)^2.$$

The first term does not depend on c and the second is never negative, so the minimum is at $c = \mu = 4.4$, where the value is $\sigma^2 = 7.84$.

This splits the expected squared miss in (b) into two pieces: $8.65 = 7.84 + 0.81$, the variance (how unpredictable inflation is) plus the squared bias $(\mu - 3.5)^2$ (how far the average outcome sits from the target). The mean is the best single-number forecast under squared loss, and the variance is exactly the loss that remains when you use it.

Problem 16

seed 20260925

A bond fund holds one bond from each of three issuers: a Georgian utility, a Kazakh telecom operator and a Turkish retailer. Over the next year the probabilities that each issuer defaults are 0.03, 0.12 and 0.08 respectively, and the fund's analysts treat the three defaults as independent events. Let Y be the number of the three bonds that default.

Give every probability to at least four significant figures.

(a) $p(0) =$ ______________

(b) $p(1) =$ ______________

(c) $P(Y \geq 2) =$ ______________

(d) $E(Y) =$ ______________

Write D_i for "issuer i defaults". The sample space has eight points, one for each pattern of default and survival, such as $D_1 \cap \bar{D}_2 \cap \bar{D}_3$. By independence, the probability of each sample point is the product of the three marginal probabilities. By Definition 3.2, $p(y)$ is the sum over the sample points with exactly y defaults.

(a) One sample point: $p(0) = (0.97)(0.88)(0.92) = 0.785312$.

(b) Three sample points, one for each issuer that could be the one to default:

$p(1) = (0.03)(0.88)(0.92) + (0.97)(0.12)(0.92) + (0.97)(0.88)(0.08) = 0.024288 + 0.107088 + 0.068288 = 0.199664.$

(c) Similarly $p(2) = 0.014736$ (three sample points) and $p(3) = (0.03)(0.12)(0.08) = 0.000288$, so $P(Y \geq 2) = 0.014736 + 0.000288 = 0.015024$. Equivalently $1 - p(0) - p(1)$, by Theorem 3.1.

(d) Definition 3.4: $E(Y) = 0 \cdot p(0) + 1 \cdot p(1) + 2 \cdot p(2) + 3 \cdot p(3) = 0.2300$.

This equals $0.03 + 0.12 + 0.08 = 0.23$, the sum of the three default probabilities. That is no accident: Y is the sum of three indicator variables, each with mean equal to its default probability, and the expectation of a sum is the sum of the expectations (Chapter 5 makes this precise for several random variables). Independence was needed to find $p(y)$ but is not needed for the mean.

Problem 17

seed 20260925

The cash-processing centre of a Baku bank is choosing between two banknote sorting machines. If a machine runs for t hours in a day, the number of jams Y that need an engineer is a random variable with **mean and variance both equal to λt**, where $\lambda_A = 0.12$ for machine A and $\lambda_B = 0.07$ for machine B. Because each extra call-out in a day is charged at a rising rate, the daily costs (AZN) are

$$C_A(t) = 4.06\,t + 40\,Y_A^2, \qquad C_B(t) = 11.00\,t + 40\,Y_B^2.$$

Assume each machine starts every day as good as new.

Give every number to at least four significant figures.

(a) For a working day of 10 hours: $E[C_A(10)] =$ ____________ AZN and $E[C_B(10)] =$ ____________ AZN

(b) Which machine has the lower expected daily cost on a day of 10 hours?

- ◯ Machine A
- ◯ Machine B

(c) Which machine has the lower expected daily cost on a day of 23 hours?

- ◯ Machine A
- ◯ Machine B

(d) The length of day $t > 0$ at which the two expected costs are equal: $t =$ ____________ hours

The cost involves Y^2, so the mean of Y alone is not enough. Theorem 3.6 rearranged gives $E(Y^2) = V(Y) + [E(Y)]^2 = \lambda t + (\lambda t)^2$. By Theorems 3.3-3.5,

$$E[C_A(t)] = 4.06t + 40\big[0.12t + (0.12t)^2\big] = 8.86\,t + 0.5760\,t^2,$$
$$E[C_B(t)] = 11.00t + 40\big[0.07t + (0.07t)^2\big] = 13.80\,t + 0.1960\,t^2.$$

(a) At $t = 10$: $E[C_A] = 146.20$ AZN and $E[C_B] = 157.60$ AZN. (For instance, for machine A, $E(Y_A) = V(Y_A) = 1.20$.)

(b) The cheaper one is Machine A.

(c) At $t = 23$: $E[C_A] = 508.48$ and $E[C_B] = 421.08$, so the cheaper one is Machine B.

(d) Set the two expressions equal and divide by $t > 0$:

$$8.86 + 0.5760\,t = 13.80 + 0.1960\,t \quad \Longrightarrow \quad t = \frac{4.94}{0.3800} = 13.0000 \text{ hours.}$$

(Both differences have the same sign, so the ratio is positive.)

Machine A is cheaper to run per hour, so it wins on days shorter than 13.0000 hours. But it jams more often, and the repair cost grows with the **square** of the number of jams, so the t^2 term eventually dominates and Machine B wins on longer days, which is exactly what (b) and (c) show. The variance enters through $E(Y^2)$: a machine with more variable jam counts is penalised even beyond its higher mean.

Problem 18

seed 20260925

A corporate-banking relationship manager in Baku meets either one client a day, with probability 0.25, or two clients, with probability 0.75. Each meeting, independently of the others, ends either with a new credit line of 50 thousand AZN, with probability 0.28, or with nothing, with probability 0.72. Let X be the total value of new credit lines arranged in a day (thousand AZN), so X takes the values 0, 50 and 100.

Give every answer to at least four significant figures.

(a) $P(X = 0) =$ ______________

(b) $P(X = 100) =$ ______________

(c) $E(X) =$ ______________ thousand AZN

(d) The standard deviation of X: ______________ thousand AZN

The sample points are the possible days: one meeting with or without a deal, or two meetings with 0, 1 or 2 deals. Each sample point's probability is the probability of the number of meetings times the probability of the deal pattern, and by Definition 3.2 each $p(x)$ adds up the sample points with that total.

(a) No deal happens either on a one-meeting day with no deal or on a two-meeting day with two failures:

$$P(X = 0) = (0.25)(0.72) + (0.75)(0.72)^2 = 0.568800.$$

(b) A total of 100 needs two meetings, both successful: $P(X = 100) = (0.75)(0.28)^2 = 0.058800$.

The remaining value has $P(X = 50) = (0.25)(0.28) + (0.75)\,2(0.28)(0.72) = 0.372400$ (on a two-meeting day the single deal can come from either meeting), and the three probabilities sum to 1 as Theorem 3.1 requires.

(c) Definition 3.4: $E(X) = 0 + 50(0.372400) + 100(0.058800) = 24.5000$.

As a check: the expected number of meetings is $1 + 0.75$, each worth 0.28×50 on average, and $50 \times 0.4900 = 24.5000$.

(d) Theorem 3.2: $E(X^2) = 50^2(0.372400) + 100^2(0.058800) = 1519.0000$, and by Theorem 3.6

$$V(X) = 1519.0000 - 24.5000^2 = 918.7500, \qquad \sigma = 30.3109.$$

The standard deviation is larger than the mean: daily deal flow is lumpy, and most days bring nothing. Here $\mu + 2\sigma = 85.1218$, which lies above 50, so even a single-deal day is within two standard deviations of the mean and does not count as unusual, although it is well above the mean. For a distribution this skewed, "within two standard deviations" is a rough guide at best; Tchebysheff's theorem (Section 3.11) only guarantees that at least 75% of the probability lies there.

Problem 19

seed 20260925

A bank tracks how the internal credit ratings of its corporate borrowers move over a year. For a randomly chosen borrower let Y be the change in rating, in notches (positive for an upgrade, negative for a downgrade). The bank's model is

y	-2	-1	0	1	2
$p(y)$	0.04	a	0.59	b	0.04

where a and b were lost when the report was reformatted. The report still states that the mean rating change is $E(Y) = -0.05$ notches.

Give every answer to at least four significant figures.

(a) $a = P(Y = -1) =$ ______________

(b) $b = P(Y = 1) =$ ______________

(c) $V(Y) =$ ______________

Two unknowns need two equations, and the section provides exactly two facts about any probability distribution with a known mean.

Theorem 3.1 (probabilities sum to 1):

$$0.04 + a + 0.59 + b + 0.04 = 1 \implies a + b = 0.33.$$

Definition 3.4 (the stated mean):

$$-2(0.04) - a + 0 + b + 2(0.04) = -0.05 \implies b - a = -0.05 + 0.08 - 0.08 = -0.05.$$

(a), (b) Adding and subtracting the two equations:

$$b = \frac{0.33 - 0.05}{2} = 0.14, \qquad a = \frac{0.33 + 0.05}{2} = 0.19.$$

Both lie in $[0, 1]$, so the completed table satisfies Theorem 3.1.

(c) Theorem 3.2 with $g(Y) = Y^2$:

$$E(Y^2) = 4(0.04) + 1(0.19) + 0 + 1(0.14) + 4(0.04) = 0.65,$$

and Theorem 3.6 gives $V(Y) = 0.65 - (-0.05)^2 = 0.6475$.

Notice that the mean alone could never have told the risk team how often ratings move: a mean near zero is consistent with almost no migration or with large offsetting moves. The variance is what measures how much ratings churn.

Problem 20

seed 20260925

An insurer has written a large property policy on a shopping centre in Baku. Its loss Y on this policy over the year (thousand AZN) is modelled as

y	0	25	150	275
$p(y)$	0.77	0.15	0.05	0.03

To cap its exposure it buys an **excess-of-loss** reinsurance contract with a retention of 75 thousand AZN: the reinsurer pays whatever part of the loss exceeds 75, that is

$$g(Y) = \max(Y - 75,\ 0).$$

Give every answer to at least four significant figures.

(a) The insurer's expected loss before reinsurance, $E(Y) =$ ______________ thousand AZN

(b) The reinsurer's expected payout, $E[g(Y)] =$ ______________ thousand AZN

(c) The standard deviation of the reinsurer's payout: ______________ thousand AZN

(d) The reinsurer charges its expected payout plus a safety loading of 0.35 times the expected payout. The premium is ______________ thousand AZN

(a) Definition 3.4: $E(Y) = 0(0.77) + 25(0.15) + 150(0.05) + 275(0.03) = 19.50$.

(b) First tabulate $g(y)$ at each possible loss. Losses of 0 and 25 are below the retention, so the reinsurer pays nothing on them:

y	0	25	150	275
$g(y)$	0	0	75	200
$p(y)$	0.77	0.15	0.05	0.03

Then Theorem 3.2:

$$E[g(Y)] = \sum_y g(y)\,p(y) = 0 + 0 + 75(0.05) + 200(0.03) = 9.75.$$

(c) Theorem 3.2 again, with $g(y)^2$: $E[g(Y)^2] = 75^2(0.05) + 200^2(0.03) = 1481.25$. By Theorem 3.6,

$$V[g(Y)] = 1481.25 - 9.75^2 = 1386.1875, \qquad \sigma = 37.2315.$$

(d) By Theorem 3.4 the premium is $E[g(Y)] + 0.35\,E[g(Y)] = 1.35 \times 9.75 = 13.1625$.

A shortcut that fails: since $E(Y) = 19.50 < 75$, plugging the mean into g gives $g(E(Y)) = \max(19.50 - 75, 0) = 0$, a contract worth nothing. The contract has value precisely because of the losses in the tail: with probability 0.92 it pays nothing, and the rest of the time it pays a lot. The standard deviation, several times the mean payout, is why the reinsurer insists on a loading.

Answer key

Problem	Answers
1	(a) 0.6, (b) 0.2, (c) 0.37
2	(a) 0.9, (b) 0.05, (c) 0.58
3	(a) 0.0217391, (b) 0.543478, (c) 2.3913, (d) 21.5217
4	(a) 0.25, (b) 550, (c) 0.17
5	(a) 258.5, (b) 353.5, (c) 6.5
6	(a) 700.8, (b) 231987, (c) 481.65
7	(a) 0.85, (b) 1.87, (c) 1.1475, (d) 1.07121
8	(a) 2.57, (b) 149.508, (c) 148.722

9	(a) 0.509091, (b) 0.436364, (c) 0.545455, (d) 245.455
10	(a) 2.8, (b) 1.28, (c) 22.4, (d) 6.22254
11	(a) 7.29, (b) 8.53, (c) 5.71541, (d) 5.25634, (e) Project ... deviation
12	(a) 3.45, (b) 1.47224, (c) 0.71, (d) 0.96
13	(a) 901, (b) 1081, (c) 7529.89
14	(a) 0.78, (b) 1.48, (c) 2119
15	(a) 27.2, (b) 8.65, (c) 4.4, (d) 7.84
16	(a) 0.785312, (b) 0.199664, (c) 0.015024, (d) 0.23
17	(a) 146.2, (b) 157.6, (c) Machine A, (d) Machine B, (e) 13
18	(a) 0.5688, (b) 0.0588, (c) 24.5, (d) 30.3109
19	(a) 0.19, (b) 0.14, (c) 0.6475
20	(a) 19.5, (b) 9.75, (c) 37.2315, (d) 13.1625