Week 8, Problem Set 2 - worked solutions

Wackerly 3.10-3.11 - Probability-Generating Functions, Tchebysheff's Theorem, and Chapter 3 in Review

> 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 bank's Baku branch network has recorded, over several years, the weekly number Y of mortgage applications it receives. The long-run mean is $\mu = 156$ and the standard deviation is $\sigma = 24$; nothing else is known about the shape of the distribution.

The mortgage desk wants to know how likely it is that next week's count falls strictly between 96 and 216.

Give every probability to at least four significant figures.

(a) Writing the interval as $(\mu - k\sigma,\ \mu + k\sigma)$, the value of k is ____________

(b) The best lower bound that Tchebysheff's theorem gives for $P(96 < Y < 216)$ is ____________

(c) The corresponding upper bound for the probability that next week's count is 96 or fewer, or 216 or more, is ____________

(d) Suppose instead that the standard deviation were only $\sigma = 12$ (same mean). The Tchebysheff lower bound for $P(96 < Y < 216)$ would then be ____________

Theorem 3.14 needs only the mean and a finite variance, which is exactly what the desk has. It says that for any $k > 0$

$$P(|Y - \mu| < k\sigma) \geq 1 - \frac{1}{k^2}, \qquad P(|Y - \mu| \geq k\sigma) \leq \frac{1}{k^2}.$$

(a) The interval is centred at the mean: $156 - 96 = 216 - 156 = 60$. Its half-width in standard deviations is $k = 60/24 = 2.5$.

(b) $P(96 < Y < 216) = P(|Y - 156| < 2.5\sigma) \geq 1 - 1/6.25 = 0.8400$.

(c) "96 or fewer, or 216 or more" is the complementary event $|Y - \mu| \geq k\sigma$, so its probability is at most $1/k^2 = 1/6.25 = 0.1600$. The two bounds in (b) and (c) are the two forms of the same statement.

(d) With $\sigma = 12$ the same interval is $k = 60/12 = 5$ standard deviations wide on each side, so the bound becomes $1 - 1/25.00 = 0.9600$.

As in Example 3.28, halving σ doubles k and pushes the guaranteed probability much closer to one. The bound is conservative: the true probability is usually well above it, but no distribution with this mean and standard deviation can fall below it.

Problem 2

seed 20260925

An insurance agency in Ganja records Y, the number of motor-insurance claims it receives in a month. From past data $E(Y) = 64$ and $V(Y) = 144$. The shape of the distribution is not known.

Give every answer to at least four significant figures.

(a) Using Tchebysheff's theorem, a lower bound for $P(46 < Y < 82)$ is ____________

(b) The agency wants a margin C such that, whatever the distribution, $P(|Y - 64| \geq C) \leq 0.18$. The smallest C that Tchebysheff's theorem guarantees this for is $C =$ ____________

(c) The upper end of the interval $\mu \pm C$ found in (b) is ____________

Here $\mu = 64$ and $\sigma = \sqrt{144} = 12$.

(a) The interval is $\mu \pm 18$, so $k\sigma = 18$ and $k = 18/12 = 1.5000$. Theorem 3.14 gives

$$P(46 < Y < 82) = P(|Y - \mu| < k\sigma) \geq 1 - \frac{1}{k^2} = 1 - \frac{\sigma^2}{(k\sigma)^2} = 1 - \frac{144}{324} = 0.5556.$$

(b) Write $C = k\sigma$. Theorem 3.14 gives $P(|Y - \mu| \geq k\sigma) \leq 1/k^2$, so the guarantee holds as soon as $1/k^2 = 0.18$, that is $k = 1/\sqrt{0.18}$ and

$$C = \frac{\sigma}{\sqrt{0.18}} = \frac{12}{0.4243} = 28.2843.$$

Any smaller C would make $1/k^2$ exceed 0.18, and Tchebysheff would no longer promise anything that strong.

(c) $\mu + C = 64 + 28.2843 = 92.2843$. Months with counts at or beyond $\mu \pm C$ together have probability at most 0.18.

Problem 3

seed 20260925

A microfinance lender in Lankaran tracks Y, the number of late instalments a borrower makes on a small business loan over one quarter (three instalments). Across its book the distribution is

y	0	1	2	3
$p(y)$	0.71	0.13	0.14	0.02

Give every number to at least four significant figures.

(a) The probability-generating function of Y is $P(t) =$ ______________

(Enter a function of t, for example 0.5 + 0.3*t + 0.2*t^2.)

(b) The first factorial moment $\mu_{[1]} = P^{(1)}(1) =$ ______________

(c) The second factorial moment $\mu_{[2]} = E[Y(Y-1)] =$ ______________

(d) $V(Y) =$ ______________

(e) The third factorial moment $\mu_{[3]} = E[Y(Y-1)(Y-2)] =$ ______________

(a) Definition 3.15: $P(t) = E(t^Y) = \sum_i p_i t^i$, so the table's probabilities become the coefficients:

$$P(t) = 0.71 + 0.13t + 0.14t^2 + 0.02t^3.$$

Check: $P(1) = 0.71 + 0.13 + 0.14 + 0.02 = 1$, as it must be for any probability-generating function.

(b) Differentiating, $P^{(1)}(t) = 0.13 + 2(0.14)t + 3(0.02)t^2$. By Theorem 3.13,

$$\mu_{[1]} = P^{(1)}(1) = 0.13 + 2(0.14) + 3(0.02) = 0.4700 = E(Y).$$

(c) $P^{(2)}(t) = 2(0.14) + 6(0.02)t$, so $\mu_{[2]} = P^{(2)}(1) = 2(0.14) + 6(0.02) = 0.4000$.

(d) Since $E[Y(Y-1)] = E(Y^2) - E(Y)$, we have $E(Y^2) = \mu_{[2]} + \mu_{[1]}$, and Theorem 3.6 gives

$$V(Y) = \mu_{[2]} + \mu_{[1]} - \mu_{[1]}^2 = 0.4000 + 0.4700 - 0.220900 = 0.6491.$$

(e) $P^{(3)}(t) = 6(0.02)$, a constant, so $\mu_{[3]} = 0.1200$. This is also clear directly: $Y(Y-1)(Y-2)$ is zero unless $Y = 3$, where it equals 6.

Problem 4

seed 20260925

A bank has just issued 38 small-business loans to unrelated firms in different regions. Each loan, independently of the others, defaults within a year with probability 0.15. Let Y be the number of these loans that default within a year.

Give every number to at least four significant figures.

(a) Which probability distribution does Y have?

- ◯ Binomial
- ◯ Geometric
- ◯ Negative binomial
- ◯ Hypergeometric
- ◯ Poisson

(b) $E(Y) =$ ______________

(c) $V(Y) =$ ______________

(a) Binomial. The situation has a fixed number $n = 38$ of independent trials, each with the same probability of success (default) $p = 0.15$, and Y counts the successes.

The five Chapter 3 models are distinguished by three questions: is the number of trials fixed in advance (binomial, hypergeometric) or is the experiment run until some number of successes (geometric, negative binomial)? Are the trials independent with a constant probability (binomial) or is the sample drawn without replacement from a small finite population (hypergeometric)? Or is there no natural “trial” at all, only events arriving at a rate (Poisson)?

(b) $E(Y) = np = 38 \times 0.15 = 5.7000$.

(c) $V(Y) = npq = 38 \times 0.15 \times 0.85 = 4.8450$.

Problem 5

seed 20260925

A bank’s call centre makes 160 outbound calls offering a new credit card. Each call reaches a different customer, and each customer independently accepts the offer with probability 0.28. Let Y be the number of acceptances.

Give every number to at least four significant figures.

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

(b) $\sigma =$ ______________

(c) The interval $\mu \pm 2\sigma$ runs from ______________ to ______________. (By Tchebysheff’s theorem it contains Y with probability at least $3/4$.)

(d) The product team says the campaign is a success only if at least 65 customers accept. Use Tchebysheff’s theorem to give an upper bound for $P(Y \geq 65)$: ______________

Y is binomial with $n = 160$ and $p = 0.28$ (fixed number of independent calls, constant acceptance probability).

(a) Theorem 3.7: $\mu = np = 160(0.28) = 44.80$.

(b) $\sigma^2 = npq = 160(0.28)(0.72) = 32.2560$, so $\sigma = 5.6794$.

(c) $\mu - 2\sigma = 33.4411$ and $\mu + 2\sigma = 56.1589$.

(d) The target sits $65 - 44.80 = 20.20$ above the mean, which is $k = 20.20/5.6794 = 3.5567$ standard deviations. Reaching it is contained in the event $|Y - \mu| \geq k\sigma$, so by Theorem 3.14

$$P(Y \geq 65) \leq P(|Y - \mu| \geq k\sigma) \leq \frac{1}{k^2} = \frac{\sigma^2}{(65-\mu)^2} = \frac{32.2560}{20.20^2} = 0.0791.$$

So success is unlikely, and this holds whatever the shape of the distribution. Because the bound also covers the lower tail, the true probability is smaller still; the binomial is close to bell-shaped here and the exact value is far below this bound.

Problem 6

seed 20260925

The number Y of cross-border payment messages that fail validation at a Baku bank in one day has a Poisson distribution with mean $\lambda = 3.18$.

Give every number to at least four significant figures.

(a) The probability-generating function of Y is $P(t) =$ ______________

(Enter a function of t; e^(...) and exp(...) both work.)

(b) $\mu_{[1]} = P^{(1)}(1) =$ ______________

(c) $\mu_{[2]} = P^{(2)}(1) =$ ______________

(d) $V(Y) =$ ______________

(e) Each failed message is repaired automatically, independently of the others, with probability 0.96; the rest need a person. The probability that **every** failed message today is repaired automatically (which includes the case of no failures at all) equals $P(0.96)$. Find it: ______________

(a) Definition 3.15 with the Poisson probability function:

$$P(t) = E(t^Y) = \sum_{y=0}^{\infty} t^y \frac{\lambda^y e^{-\lambda}}{y!} = e^{-\lambda} \sum_{y=0}^{\infty} \frac{(\lambda t)^y}{y!} = e^{-\lambda} e^{\lambda t} = e^{\lambda(t-1)} = e^{3.18(t-1)},$$

finite for every t.

(b) $P^{(1)}(t) = \lambda e^{\lambda(t-1)}$, so by Theorem 3.13 $\mu_{[1]} = P^{(1)}(1) = \lambda = 3.18 = E(Y)$.

(c) $P^{(2)}(t) = \lambda^2 e^{\lambda(t-1)}$, so $\mu_{[2]} = E[Y(Y-1)] = \lambda^2 = 10.1124$.

(d) $E(Y^2) = \mu_{[2]} + \mu_{[1]}$, hence $V(Y) = \lambda^2 + \lambda - \lambda^2 = \lambda = 3.1800$, the familiar Poisson result, obtained here without summing a single series by hand.

(e) Given $Y = y$, all y failures are repaired automatically with probability s^y (independence), and with probability 1 when $y = 0$. Averaging over y (law of total probability),

$$P(\text{all repaired}) = \sum_y s^y p(y) = E(s^Y) = P(s) = e^{3.18(-0.04)} = e^{-0.1272} = 0.8806.$$

This is one of the uses of $P(t)$ that the moment-generating function does not have: for $0 \leq s \leq 1$, $P(s)$ is itself a probability.

Problem 7

seed 20260925

A Baku brokerage opens, on average, $\mu = 200$ new retail trading accounts per business day, with standard deviation $\sigma = 14$. The daily count is clearly not bell-shaped (it jumps on days with IPO news), so the compliance team will rely only on Tchebysheff's theorem.

Give every number to at least four significant figures.

(a) The smallest k for which Tchebysheff's theorem guarantees $P(|Y-\mu|<k\sigma) \geq 0.88$ is $k=$ ______________

(b) The corresponding interval $(\mu-k\sigma,\ \mu+k\sigma)$, which contains the daily count with probability at least 0.88, has lower end ______________ and upper end ______________

(c) The team also wants a half-width w with $P(|Y-\mu| \geq w) \leq 0.09$ guaranteed. The smallest such w is ______________

(a) Theorem 3.14 promises $P(|Y-\mu|<k\sigma) \geq 1-1/k^2$. The promise reaches 0.88 exactly when

$$1-\frac{1}{k^2}=0.88 \iff k^2=\frac{1}{0.12} \iff k=\frac{1}{\sqrt{0.12}}=2.8868.$$

A smaller k gives a smaller bound, so this is the smallest k that works.

(b) $k\sigma = 2.8868 \times 14 = 40.4145$, so the interval is 200 ± 40.4145, from 159.5855 to 240.4145.

(c) Now it is the tail form: $P(|Y-\mu| \geq k\sigma) \leq 1/k^2 = 0.09$ requires $k = 1/\sqrt{0.09} = 3.3333$, so $w = k\sigma = 3.3333 \times 14 = 46.6667$.

Notice how quickly the guaranteed interval widens as the target approaches 1: Tchebysheff's k grows like $1/\sqrt{1-f}$, whereas for a bell-shaped distribution the empirical rule would already give about 95% at $k=2$. That is the price of assuming nothing about the shape.

Problem 8

seed 20260925

The number of household-insurance claims lodged per day with a large insurer follows a Poisson distribution with mean $\lambda = 108$.

The claims department wants an interval, centred at the mean, that is guaranteed by Tchebysheff's theorem to contain the daily count on at least $24/25$ of all days.

Give every number to at least four significant figures.

(a) The standard deviation of the daily count is $\sigma =$ ______________

(b) The interval runs from ______________ to ______________

(c) Over a review period of 264 working days, the expected number of days on which the count falls inside this interval is at least

____

(a) For a Poisson random variable $E(Y) = V(Y) = \lambda$ (Theorem 3.11), so $\sigma = \sqrt{108} = 10.3923$.

(b) We need $1 - 1/k^2 = 24/25$, so $1/k^2 = 1/25$, $k^2 = 25.00$ and $k = 5$. Theorem 3.14 then gives

$$P(\lambda - 5\sigma < Y < \lambda + 5\sigma) \geq 1 - \frac{1}{25.00} = \frac{24}{25},$$

with $k\sigma = 5 \times 10.3923 = 51.9615$. The interval is 108 ± 51.9615, that is from 56.0385 to 159.9615.

(c) Let X be the number of the 264 days with the count inside the interval. Each day is inside with probability $p \geq 24/25$, so $E(X) = 264p \geq 264 \times 24/25 = 253.4400$.

Tchebysheff did not use the Poisson shape at all, only σ. For a Poisson with mean this large the distribution is close to bell-shaped, and the actual fraction of days inside the interval is well above $24/25$.

Problem 9

seed 20260925

A loan officer reviews 4 credit applications from unrelated applicants this morning. Each application, independently of the others, is approved with probability 0.555. Let Y be the number approved.

Give every number to at least four significant figures.

(a) $\mu =$ ______________ and $\sigma =$ ______________

(b) Using the binomial probability function, the exact probability $P(|Y-\mu| < \sigma) =$ ______________

(c) The exact probability $P(|Y-\mu| < 2\sigma) =$ ______________

(d) Tchebysheff's theorem guarantees $P(|Y-\mu| < 2\sigma) \geq 0.75$. By how much does the exact probability in (c) exceed this bound?

____

Y is binomial with $n = 4$, $p = 0.555$, $q = 0.445$.

(a) $\mu = np = 2.220$ and $\sigma = \sqrt{npq} = 0.9939$.

The probability function $p(y) = \binom{4}{y}(0.555)^y(0.445)^{4-y}$ gives

y	0	1	2	3	4
$p(y)$	0.0392	0.1956	0.3660	0.3043	0.0949

(b) $\mu \pm \sigma$ is the interval from 1.2261 to 3.2139. The values of y strictly inside it are $2, 3$, so

$$P(|Y-\mu| < \sigma) = 0.6703.$$

(c) $\mu \pm 2\sigma$ runs from 0.2321 to 4.2079. The values inside are $1, 2, 3, 4$, so $P(|Y-\mu| < 2\sigma) = 0.9608$.

(d) $0.9608 - 0.75 = 0.2108$.

Tchebysheff's bound holds, as it must, but it is far from tight: for $k = 1$ it says nothing at all ($1 - 1/1^2 = 0$), and for $k = 2$ the exact probability is well above $3/4$. Compare the empirical rule's approximately 68% and 95%: with only 4 trials the distribution is lumpy, and when p is far from $1/2$ it is skewed, so the empirical rule can be off in either direction, whereas Tchebysheff's bound can never fail.

Problem 10

seed 20260925

The container terminal at the Port of Baku (Alat) handles Y containers per day, with mean $\mu = 380$ and standard deviation $\sigma = 38$. Nothing else is assumed about the distribution.

The terminal's staffing plan works well when $285 < Y < 532$.

Give every probability to at least four significant figures.

(a) The staffing interval is not centred at μ. The largest interval of the form $(\mu - k\sigma,\ \mu + k\sigma)$ that fits inside it has $k =$

____

(b) Hence Tchebysheff's theorem gives the lower bound $P(285 < Y < 532) \geq$ ______________

(c) An upper bound for $P(Y \geq 532)$ from Theorem 3.14 is ______________

(d) An upper bound for $P(Y \leq 285)$ from Theorem 3.14 is ______________

The lower end is $380 - 285 = 95 = 2.5\sigma$ below the mean and the upper end is $532 - 380 = 152 = 4\sigma$ above it.

(a) Theorem 3.14 speaks only about intervals centred at μ. A centred interval fits inside $(285, 532)$ only if its half-width does not exceed the distance to the nearer end, here the lower end, so the largest one has $k = \min(2.5, 4) = 2.5$: it is $(285, 475)$.

(b) Because $(285, 475) \subseteq (285, 532)$,

$$P(285 < Y < 532) \geq P(|Y - \mu| < 2.5\sigma) \geq 1 - \frac{1}{6.25} = 0.8400.$$

Using the farther end instead would claim more than the theorem delivers.

(c) $Y \geq 532$ means $Y - \mu \geq 4\sigma$, which is contained in $|Y - \mu| \geq 4\sigma$. So $P(Y \geq 532) \leq 1/16.00 = 0.0625$.

(d) Likewise $Y \leq 285$ is contained in $|Y - \mu| \geq 2.5\sigma$, so $P(Y \leq 285) \leq 1/6.25 = 0.1600$.

Adding (c) and (d) gives another valid bound for the staffing interval, $1 - 0.0625 - 0.1600$, but it is always weaker than (b): it subtracts the two tails separately, each with its own $1/k^2$, whereas the single bound in (b) covers both tails beyond 2.5σ at once.

Problem 11

seed 20260925

A card-payment processor in Baku declines, on average, $\mu = 800$ transactions a day. Its service agreement with a partner bank asks that the daily number Y of declines lie strictly between 655 and 945 (that is, within ± 145 of the mean) with probability at least 0.90.

The distribution of Y is unknown, so the processor wants the agreement to be guaranteed by Tchebysheff's theorem alone.

Give every number to at least four significant figures.

(a) The largest standard deviation σ for which Tchebysheff's theorem guarantees $P(|Y - \mu| < 145) \geq 0.90$ is ______________

(b) In fact the standard deviation is $\sigma = 73.95$. The lower bound that Tchebysheff's theorem gives for $P(|Y - \mu| < 145)$ is

____

(c) With $\sigma = 73.95$, does Tchebysheff's theorem alone guarantee the agreement?

○ Yes
○ No

(d) With $\sigma = 73.95$, the narrowest band $\mu \pm w$ that Tchebysheff's theorem guarantees to hold with probability at least 0.90 has $w =$

____

Write the band as $145 = k\sigma$. Theorem 3.14 gives $P(|Y - \mu| < 145) \geq 1 - 1/k^2 = 1 - \sigma^2/145^2$.

(a) We need $1 - \sigma^2/21025 \geq 0.90$, i.e. $\sigma^2 \leq 0.10 \times 21025$, so

$$\sigma \leq 145\sqrt{0.10} = 145 \times 0.3162 = 45.8530.$$

(b) $1 - 5468.6025/21025 = 0.7399$.

(c) No: $\sigma = 73.95 > 45.8530$, so the bound in (b) is 0.7399 against the required 0.90. This does not mean the agreement will fail; it means Tchebysheff cannot promise it, and the processor would need to know more about the shape of the distribution.

(d) Solve $1 - \sigma^2/w^2 = 0.90$ for w: $w = \sigma/\sqrt{0.10} = 73.95/0.3162 = 233.8504$. Answers (a) and (d) are the same relation $w = \sigma/\sqrt{1-f}$ read in two directions: fix the band and find the largest σ, or fix σ and find the narrowest band.

Problem 12

seed 20260925

A Baku freight forwarder insures 9 export shipments this month, each with a different client on a different route. Each shipment, independently of the others, produces an insurance claim with probability 0.27. Let Y be the number of claims.

Give every number to at least four significant figures.

(a) The probability-generating function of Y is $P(t) =$ ______________

(Enter a function of t.)

(b) Using Theorem 3.13, $E(Y) = P^{(1)}(1) =$ ______________

(c) $\mu_{[2]} = E[Y(Y-1)] =$ ______________

(d) $V(Y) =$ ______________

(e) Evaluate $P^{(1)}(0)$; it equals $P(Y = 1)$. ______________

(f) Each claim, independently, is settled without litigation with probability 0.80. The probability that no claim this month goes to court (including the case of no claims) is $P(0.80)$: ______________

(a) By Definition 3.15 and the binomial theorem,

$$P(t) = \sum_{y=0}^{9} t^y \binom{9}{y} p^y q^{9-y} = \sum_{y=0}^{9} \binom{9}{y} (pt)^y q^{9-y} = (q + pt)^9 = (0.73 + 0.27t)^9.$$

(b) $P^{(1)}(t) = 9p(q + pt)^8$. At $t = 1$, $q + p = 1$, so $E(Y) = 9p = 9(0.27) = 2.4300$.

(c) $P^{(2)}(t) = 9(8)p^2(q + pt)^7$, so $\mu_{[2]} = 9(8)(0.27)^2 = 5.2488$.

(d) $V(Y) = \mu_{[2]} + \mu_{[1]} - \mu_{[1]}^2 = 5.2488 + 2.4300 - (2.4300)^2 = 1.7739$, which is npq, as Theorem 3.7 says.

(e) $P^{(1)}(0) = 9p\,q^8 = 9(0.27)(0.73)^8 = 0.1960$. Differentiating $P(t) = p_0 + p_1 t + p_2 t^2 + \cdots$ once and setting $t = 0$ leaves exactly p_1: the coefficients of $P(t)$ are the probabilities.

(f) Given $Y = y$, all y claims avoid court with probability s^y, so the probability is $E(s^Y) = P(s)$:

$$P(0.80) = (0.73 + 0.27 \times 0.80)^9 = (0.9460)^9 = 0.6068.$$

Problem 13

seed 20260925

A bank models the month Y in which a newly issued credit card first has a missed payment. Each month, independently, the cardholder misses a payment with probability $p = 0.066$, so Y is geometric: $p(y) = q^{y-1}p$, $y = 1, 2, \ldots$, with $q = 1 - p$.

Example 3.26 derived its probability-generating function, $P(t) = \dfrac{pt}{1 - qt}$, valid for $t < 1/q$.

Give every number to at least four significant figures.

(a) $P(t)$ is finite for t below ______________

(b) $\mu_{[1]} = P^{(1)}(1) =$ ______________

(c) $\mu_{[2]} = P^{(2)}(1) =$ ______________

(d) $V(Y) =$ ______________

(e) When the first miss happens the bank books a provision of 2500 AZN in that month. With a monthly discount factor $v = 0.999$, the present value of the provision is $2500\,v^Y$. Its expected present value, $2500\,E(v^Y) = 2500\,P(v)$, is ______________ AZN

(a) The series $\sum_y t^y q^{y-1} p = (p/q) \sum_y (qt)^y$ is geometric and converges only when $qt < 1$, so $t < 1/q = 1/0.934 = 1.0707$. Since this is above 1, $P(t)$ and its derivatives exist at $t = 1$, which is all Theorem 3.13 needs.

(b) As in Example 3.27, the quotient rule gives $P^{(1)}(t) = \dfrac{(1-qt)p + ptq}{(1-qt)^2} = \dfrac{p}{(1-qt)^2}$, so

$$\mu_{[1]} = P^{(1)}(1) = \frac{p}{p^2} = \frac{1}{p} = \frac{1}{0.066} = 15.1515.$$

(c) Differentiating $p(1-qt)^{-2}$ again, $P^{(2)}(t) = \dfrac{2pq}{(1-qt)^3}$, and

$$\mu_{[2]} = E[Y(Y-1)] = \frac{2pq}{p^3} = \frac{2q}{p^2} = 428.8338.$$

(d) $V(Y) = \mu_{[2]} + \mu_{[1]} - \mu_{[1]}^2 = \dfrac{2q}{p^2} + \dfrac{1}{p} - \dfrac{1}{p^2} = \dfrac{2q+p-1}{p^2} = \dfrac{q}{p^2} = 214.4169$, matching Theorem 3.8.

(e) $E(v^Y)$ is the probability-generating function evaluated at $t = v$:

$$P(v) = \frac{pv}{1-qv} = \frac{0.065934}{0.066934} = 0.9851,$$

so the expected present value is $2500 \times 0.9851 = 2462.6498$ AZN. In finance $E(v^Y)$ is the expected discount factor of a payment made at a random time, and for an integer-valued time it is exactly the probability-generating function.

Problem 14

seed 20260925

An internal auditor at a bank branch pulls 8 of the branch's 48 mortgage files at random, without replacement. Unknown to the auditor, exactly 16 of the 48 files have a documentation error. Let Y be the number of flawed files among those pulled.

Give every number to at least four significant figures, as a decimal.

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

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

(c) $E(Y) =$ ______________

(d) $V(Y) =$ ______________

(e) Had the auditor put each file back before drawing the next (sampling with replacement), Y would be binomial. Its variance would then be ______________

Drawing without replacement from a finite population with $r = 16$ "successes" among $N = 48$ makes Y hypergeometric: $p(y) = \binom{r}{y}\binom{N-r}{n-y}\big/\binom{N}{n}$.

(a) $P(Y=0) = \binom{32}{8}\big/\binom{48}{8}$. Cancelling the factorials leaves a product of 8 fractions,

$$P(Y=0) = \frac{32}{48} \cdot \frac{32-1}{48-1} \cdots \frac{32-8+1}{48-8+1} = 0.0279,$$

which is also the natural way to see it: each file pulled must be one of the clean ones that remain.

(b) $P(Y=1) = 16\binom{32}{8-1}\big/\binom{48}{8}$. Dividing by $P(Y=0)$,

$$\frac{P(Y=1)}{P(Y=0)} = \frac{16\,8}{32-8+1} \quad \Rightarrow \quad P(Y=1) = 0.1427,$$

so $P(Y \geq 2) = 1 - 0.0279 - 0.1427 = 0.8294$.

(c) Theorem 3.10: $E(Y) = nr/N = 8(16)/48 = 2.6667$.

(d) $V(Y) = n\left(\dfrac{r}{N}\right)\left(\dfrac{N-r}{N}\right)\left(\dfrac{N-n}{N-1}\right) = 8\left(\dfrac{16}{48}\right)\left(\dfrac{32}{48}\right)\left(\dfrac{40}{47}\right) = 1.5130.$

(e) With replacement, each draw is an independent trial with $p = r/N = 16/48$, so $V(Y) = np(1-p) = 1.7778$.

The two variances differ by exactly the factor $(N-n)/(N-1) = 0.8511$. Sampling without replacement is less variable, because every flawed file found leaves fewer to find. With 8 files out of 48, the correction is not negligible; it would be if the branch had thousands of files.

Problem 15

seed 20260925

A relationship manager at a Baku investment bank is placing a new corporate bond issue. She contacts institutional clients one at a time, and each contact, independently of the others, results in a placement with probability 0.48. She stops after the 2nd placement. Let Y be the number of the contact on which the 2nd placement is made.

Give every number to at least four significant figures.

(a) $P(Y=6) =$ ______________

(b) The probability that she needs at most 3 contacts, $P(Y \le 3) =$ ______________

(c) $E(Y) =$ ______________

(d) $V(Y) =$ ______________

Independent trials with constant success probability, counted until the rth success: Y has a negative binomial distribution with $r = 2$, $p = 0.48$,

$$p(y) = \binom{y-1}{r-1} p^r q^{y-r}, \qquad y = r, r+1, \dots$$

(a) The 6th contact must be a placement, and exactly 1 of the first 5 must be placements:

$$P(Y=6) = \binom{5}{1}(0.48)^2(0.52)^4 = 5(0.48)^2(0.52)^4 = 0.0842.$$

(b) Y cannot be smaller than 2, so $P(Y \le 3) = p(2) + p(3)$ with $p(2) = p^2 = 0.230400$ and $p(3) = \binom{2}{1}p^2q = 2p^2q = 0.239616$, giving 0.4700.

(c) Theorem 3.9: $E(Y) = r/p = 2/0.48 = 4.1667$.

(d) $V(Y) = rq/p^2 = 2(0.52)/(0.48)^2 = 4.5139$.

When $r = 1$ these formulas reduce to the geometric ones, $1/p$ and q/p^2; the negative binomial waiting time is a sum of r independent geometric waiting times, which is why its mean and variance are r times theirs.

Problem 16

seed 20260925

A credit-risk analyst at a Baku bank fits a model for an integer-valued count Y attached to the bank's retail loan book and reports only its moment-generating function:

$$m(t) = (0.74 + 0.26e^t)^{28}$$

Give every number to at least four significant figures.

(a) Which distribution does Y have?

- ○ Binomial
- ○ Geometric
- ○ Negative binomial

○ Hypergeometric
○ Poisson

(b) $E(Y) =$ ______________

(c) $V(Y) =$ ______________

(d) $P(Y = 2) =$ ______________

(a) Binomial, with $n = 28,\ p = 0.26$. The mgf, when it exists, determines the distribution uniquely, so matching $m(t)$ against the mgfs derived in Section 3.9 (binomial $(q + pe^t)^n$, Poisson $e^{\lambda(e^t-1)}$, geometric $pe^t/(1 - qe^t)$, negative binomial $[pe^t/(1 - qe^t)]^r$) identifies it. The hypergeometric has no simple mgf and is a distractor here.

(b)-(c) By Theorem 3.12 the moments are derivatives of $m(t)$ at $t = 0$. Here $m'(t) = 28(0.74 + 0.26e^t)^{27}(0.26e^t)$, and carrying the differentiation through gives

$$E(Y) = m'(0) = np, \quad E(Y^2) = m''(0) = n(n-1)p^2 + np, \quad V(Y) = npq,$$

so $E(Y) = 7.2800$ and $V(Y) = 5.3872$.

(d) From the probability function of the identified distribution, $p(2) = \binom{28}{2}(0.26)^2(0.74)^{26} = 0.010173$.

Problem 17

seed 20260925

A market-making desk on the Baku Stock Exchange records Y, the number of order legs it executes in a one-minute window. The risk system does not report the probability function of Y directly, only its probability-generating function

$$P(t) = e^{1.60(t^2-1)}.$$

Give every number to at least four significant figures.

(a) $P(Y = 4) =$ ______________

(b) $P(Y = 6) =$ ______________

(c) $E(Y) =$ ______________

(d) $\mu_{[2]} = E[Y(Y - 1)] =$ ______________

(e) $V(Y) =$ ______________

(f) Each leg, independently, clears without manual intervention with probability 0.90. The probability that every leg executed in the window clears automatically (including the case of no legs) is ______________

Write $\lambda = 1.60$. Expanding the exponential in powers of t,

$$P(t) = e^{-\lambda}e^{\lambda t^2} = e^{-\lambda}\sum_{j=0}^{\infty}\frac{\lambda^j t^{2j}}{j!}.$$

By Definition 3.15 the coefficient of t^y is $P(Y = y)$. Only even powers appear, so Y takes only the values $0, 2, 4, \ldots$, and $P(Y = 2j) = e^{-\lambda}\lambda^j/j!$: Y is twice a Poisson(λ) count. (Each order is executed as a matched pair of legs.)

(a) $j = 2$: $P(Y = 4) = e^{-1.60}(2.5600)/2 = 0.201897 \times 2.5600/2 = 0.2584$.

(b) $j = 3$: $P(Y = 6) = 0.201897 \times 4.0960/6 = 0.1378$.

(c) $P^{(1)}(t) = 2\lambda t\, e^{\lambda(t^2-1)}$, so by Theorem 3.13 $E(Y) = P^{(1)}(1) = 2\lambda = 3.2000$.

(d) By the product rule, $P^{(2)}(t) = (2\lambda + 4\lambda^2 t^2)e^{\lambda(t^2-1)}$, so $\mu_{[2]} = P^{(2)}(1) = 2\lambda + 4\lambda^2 = 13.4400$.

(e) $V(Y) = \mu_{[2]} + \mu_{[1]} - \mu_{[1]}^2 = 13.4400 + 3.2000 - 10.2400 = 6.4000$, which is 4λ, as it must be for $Y = 2X$ with $V(X) = \lambda$.

(f) Given $Y = y$, all legs clear with probability s^y, so the answer is $E(s^Y) = P(s) = e^{1.60(0.90^2-1)} = e^{1.60(-0.1900)} = 0.7379$.

The point of the problem is the one the book makes at the end of Section 3.10: the probability-generating function is most useful not for moments but for recovering an unfamiliar probability function from its coefficients.

Problem 18

seed 20260925

A hedged arbitrage book at a Baku bank has a daily profit-and-loss Y (thousand AZN) with only three possible values:

y	-25	0	25
$p(y)$	0.095	0.810	0.095

Take $k = 2.5$ throughout. Give every number to at least four significant figures.

(a) $E(Y) = 0$ by symmetry. The standard deviation is $\sigma =$ ______________

(b) Using the distribution itself, the exact probability $P(|Y - \mu| \geq 2.5\sigma) =$ ______________

(c) The upper bound Tchebysheff's theorem gives for this probability is ______________

(d) Keep the values $-25, 0, 25$ and the symmetry, but change the probability of each large move from 0.095 to a new value a. The value of a for which the exact probability $P(|Y - \mu| \geq 2.5\sigma)$ equals the Tchebysheff bound is $a =$ ______________

(a) $E(Y) = -25(0.095) + 0 + 25(0.095) = 0$, so $V(Y) = E(Y^2) = 2(0.095)(25)^2 = 118.7500$ and $\sigma = 10.8972$.

(b) $k\sigma = 2.5 \times 10.8972 = 27.2431$. Here $k\sigma = 27.2431 > 25$, so no possible value of Y is that far from the mean: $P(|Y| \geq k\sigma) = 0$.

(c) Theorem 3.14: $P(|Y - \mu| \geq k\sigma) \leq 1/k^2 = 1/6.25 = 0.1600$. The exact value 0.0000 respects it, as every distribution must.

(d) With probability a on each of ± 25, $\sigma = 25\sqrt{2a}$. The exact tail is $2a$ as long as $25 \geq k\sigma$, and it can equal $1/k^2$ only then. Setting $2a = 1/k^2$ gives

$$a = \frac{1}{2k^2} = \frac{1}{12.50} = 0.080000,$$

and then $\sigma = 25\sqrt{1/k^2} = 25/k = 10.0000$, so $k\sigma = 25$ exactly: the two large moves sit precisely on the edge of the Tchebysheff interval and carry all the probability the theorem allows outside it.

This is the book's point in Exercise 3.169: for any $k > 1$ some distribution attains the bound, so $1/k^2$ cannot be replaced by anything smaller without extra assumptions about the shape. The conservatism of Tchebysheff's bound is the price of protection against exactly this kind of lumpy, all-or-nothing P&L.

Problem 19

seed 20260925

A card issuer's fraud team monitors a merchant category in which 1400 transactions were made yesterday. Each transaction, independently of the others, is fraudulent with probability 0.0015. Let Y be the number of fraudulent transactions.

Give every number to at least four significant figures, as a decimal.

(a) The exact probability of at most two frauds, $P(Y \leq 2) =$ ______________

(b) The Poisson approximation to the same probability is ______________

(c) The standard deviation of Y (exact, binomial) is $\sigma =$ ______________

(d) An alarm is raised if $Y \geq 9$. Without using the shape of the distribution, Tchebysheff's theorem bounds the probability of an alarm by ______________

Y is binomial with $n = 1400$, $p = 0.0015$: a fixed number of independent transactions with constant fraud probability.

(a) $p(0) = q^{1400} = (0.9985)^{1400} = 0.122264$. Successive binomial probabilities are related by $p(y+1)/p(y) = \frac{(n-y)p}{(y+1)q}$, so

$$p(1) = 0.122264 \cdot \frac{1400\,p}{q} = 0.257139, \qquad p(2) = 0.257139 \cdot \frac{1399\,p}{2q} = 0.270209,$$

and $P(Y \leq 2) = 0.649611$.

(b) n is large and p small, so Y is approximately Poisson with $\lambda = np = 2.10$ (Section 3.8):

$$P(Y \leq 2) \approx e^{-\lambda}\left(1 + \lambda + \frac{\lambda^2}{2}\right) = 0.122456\left(1 + 2.10 + \frac{2.10^2}{2}\right) = 0.649631.$$

The two answers differ by only 0.000020, which is the whole justification for the approximation: the exact computation needs q^{1400} and the approximation needs only λ.

(c) $\sigma^2 = npq = 2.0969$, so $\sigma = 1.4481$, barely smaller than the Poisson value $\sqrt{\lambda}$ because q is almost 1.

(d) The alarm level is $9 - 2.10 = 6.90$ above the mean, that is $k = 6.90/1.4481 = 4.7650$ standard deviations. By Theorem 3.14,

$$P(Y \geq 9) \leq P(|Y - \mu| \geq k\sigma) \leq \frac{\sigma^2}{(9-\mu)^2} = \frac{2.0969}{6.90^2} = 0.0440.$$

A false alarm rate of at most this size is guaranteed by the mean and variance alone. The true rate, from the binomial itself, is far smaller: Tchebysheff pays for its generality by ignoring the shape of the distribution.

Problem 20

seed 20260925

A logistics company in Sumgayit leases a fleet of delivery vans. The number Y of breakdowns in a month is Poisson with mean $\lambda = 3.6$. The maintenance contract costs, for the month,

- a fixed fee of 6000 AZN,
- 500 AZN for every breakdown, and
- a surcharge of 400 AZN for every **pair** of breakdowns in the same month, because overlapping repairs need an extra crew.

So the monthly cost is $C = 6000 + 500Y + 400 \cdot \dfrac{Y(Y-1)}{2}$.

Give every number to at least four significant figures.

(a) Using the probability-generating function of the Poisson distribution, $\mu_{[2]} = E[Y(Y-1)] =$ ______________

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

(c) The finance team wants a monthly budget that is exceeded with probability at most $4/25$, guaranteed by Tchebysheff's theorem. First find the upper end u of the interval $\lambda \pm k\sigma$ with $1 - 1/k^2 = 21/25$: $u =$ ______________

(d) The cost at $Y = u$, the budget level, is ______________ AZN

(a) For the Poisson, $P(t) = e^{\lambda(t-1)}$ and $P^{(2)}(t) = \lambda^2 e^{\lambda(t-1)}$, so by Theorem 3.13 $\mu_{[2]} = P^{(2)}(1) = \lambda^2 = 12.9600$.

(b) The cost is linear in Y and in $Y(Y-1)$, so by Theorems 3.3-3.5

$$E(C) = 6000 + 500E(Y) + \frac{400}{2}E[Y(Y-1)] = 6000 + 500(3.6) + 200(12.9600) = 6000 + 1800.00 + 2592.00 = 10392.00.$$

The pair term is exactly why the factorial moment is the natural tool: no $E(Y^2)$ is needed.

(c) $1 - 1/k^2 = 21/25$ gives $k^2 = 6.25$, $k = 2.5$. With $\sigma = \sqrt{\lambda} = 1.8974$, $u = \lambda + k\sigma = 3.6 + 2.5(1.8974) = 8.3434$.

(d) $C(y) = 6000 + 500y + 400y(y-1)/2$ is increasing for $y \geq 0$ (its slope $500 + 400(y - 1/2)$ is positive because $400/2 < 500$). So

$$C(u) = 6000 + 500(8.3434) + 400(8.3434)(7.3434)/2 = 22425.54.$$

Because C is increasing, $C \geq C(u)$ happens only if $Y \geq u$, which is contained in $|Y - \lambda| \geq k\sigma$. Theorem 3.14 therefore gives $P(C \geq 22425.54) \leq 1/k^2 = 4/25$.

Note the budget is well above $E(C) = 10392.00$: the quadratic pair term makes the cost distribution skewed to the right, and a prudent budget has to allow for bad months, not the average one.

Answer key

Problem	Answers
1	(a) 2.5, (b) 0.84, (c) 0.16, (d) 0.96
2	(a) 0.555556, (b) 28.2843, (c) 92.2843
3	(a) 0.71+0.13*t+0.14*t^2+0.02*t^3, (b) 0.47, (c) 0.4, (d) 0.6491, (e) 0.12
4	(a) Binomial, (b) 5.7, (c) 4.845
5	(a) 44.8, (b) 5.67944, (c) 33.4411, (d) 56.1589, (e) 0.0790511
6	(a) e^[3.18*(t-1)], (b) 3.18, (c) 10.1124, (d) 3.18, (e) 0.880558
7	(a) 2.88675, (b) 159.585, (c) 240.415, (d) 46.6667
8	(a) 10.3923, (b) 56.0385, (c) 159.962, (d) 253.44
9	(a) 2.22, (b) 0.993932, (c) 0.670278, (d) 0.960786, (e) 0.210786
10	(a) 2.5, (b) 0.84, (c) 0.0625, (d) 0.16
11	(a) 45.853, (b) 0.7399, (c) No, (d) 233.85
12	(a) (0.73+0.27*t)^9, (b) 2.43, (c) 5.2488, (d) 1.7739, (e) 0.19597, (f) 0.606765
13	(a) 1.07066, (b) 15.1515, (c) 428.834, (d) 214.417, (e) 2462.65
14	(a) 0.0278742, (b) 0.82941, (c) 2.66667, (d) 1.513, (e) 1.77778
15	(a) 0.0842298, (b) 0.470016, (c) 4.16667, (d) 4.51389
16	(a) Binomial, (b) 7.28, (c) 5.3872, (d) 0.0101733
17	(a) 0.258428, (b) 0.137828, (c) 3.2, (d) 13.44, (e) 6.4, (f) 0.737861
18	(a) 10.8972, (b) 0, (c) 0.16, (d) 0.08
19	(a) 0.649611, (b) 0.649631, (c) 1.44805, (d) 0.0440422
20	(a) 12.96, (b) 10392, (c) 8.34342, (d) 22425.5