Week 6, Problem Set 1 - worked solutions

Wackerly 3.4 - The Binomial Probability Distribution

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

Problem 1

seed 20260925

A bank's card department offers a free upgrade to a premium card to 8 existing customers, chosen at random from its customer base. Each customer accepts with probability 0.38, independently of the others. Let Y be the number of the 8 customers who accept.

Give every probability to at least four significant figures.

(a) $P(Y = 3) =$ ______________

(b) What is the probability that nobody accepts? ______________

Check Definition 3.6: there is a fixed number $n = 8$ of trials (the customers); each ends in "accepts" (the success, S) or "declines"; the success probability $p = 0.38$ is the same for every customer; the customers decide independently; and Y counts the successes. So Y is binomial with $n = 8$, $p = 0.38$ and $q = 1 - p = 0.62$.

(a) By Definition 3.7, $p(y) = \binom{n}{y} p^y q^{n-y}$. A sample point with exactly 3 acceptances has probability $p^3 q^5$, and there are $\binom{8}{3} = 56$ such sample points (Theorem 2.3). Hence

$$P(Y = 3) = 56 \times 0.38^3 \times 0.62^5 = 56 \times 0.054872 \times 0.0916133 = 0.2815.$$

(b) "Nobody accepts" is the single sample point $FF \cdots F$, so

$$P(Y = 0) = \binom{8}{0} p^0 q^8 = 0.62^8 = 0.0218.$$

Problem 2

seed 20260925

A courier company in Baku insures its fleet of 14 delivery vans with one insurer. The vans work separate routes, and the insurer's actuary models each van as filing at least one claim during the policy year with probability 0.32, independently of the other vans. Let Y be the number of vans that file a claim.

Give each answer to at least four significant figures.

(a) $E(Y) =$ ______________

(b) $V(Y) =$ ______________

(c) The standard deviation of Y is ______________

(d) The expected number of vans that file **no** claim is ______________

Y counts successes ("files a claim") in $n = 14$ independent trials with the same success probability, so Y is binomial with $n = 14$ and $p = 0.32$, $q = 0.68$. Theorem 3.7 gives the moments without summing $y\,p(y)$ over all y.

(a) $E(Y) = np = 14 \times 0.32 = 4.48$.

(b) $V(Y) = npq = 14 \times 0.32 \times 0.68 = 3.0464$.

(c) $\sigma = \sqrt{npq} = \sqrt{3.0464} = 1.7454$.

(d) The vans with no claim number $14 - Y$. By Theorem 3.4 (expectation of a linear function), $E(14 - Y) = 14 - E(Y) = 14 - 4.48 = 9.52$. Equivalently, relabel "files no claim" as the success: that count is binomial with success probability q, so its mean is $nq = 9.52$, and its variance nqp is the same as in (b).

Problem 3

seed 20260925

An external auditor is checking the invoice file of a large wholesale distributor, which contains tens of thousands of invoices. Past work suggests that 15.0% of the invoices carry a pricing error. The auditor selects 10 invoices at random. Let Y be the number of sampled invoices with a pricing error; the file is so large that Y may be treated as binomial.

Give every probability to at least four significant figures.

(a) What is the probability that the sample contains at least one invoice with an error? ________________

(b) What is the probability that it contains at least two? ________________

As in Example 3.7, removing a few invoices from a very large file barely changes the error rate among those left, so Y is approximately binomial with $n = 10$, $p = 0.150$, $q = 0.850$.

"At least one" contains every value $1, 2, \ldots, 10$, while its complement contains only $y = 0$, so it is far quicker to go through the complement.

(a)

$$P(Y \geq 1) = 1 - p(0) = 1 - \binom{10}{0} p^0 q^{10} = 1 - 0.850^{10} = 1 - 0.19687 = 0.80313.$$

(b) The complement of $Y \geq 2$ is $\{Y = 0\} \cup \{Y = 1\}$, two mutually exclusive events, with $p(1) = \binom{10}{1} p\, q^9 = 10 \times 0.150 \times 0.850^9 = 0.34743$. So

$$P(Y \geq 2) = 1 - p(0) - p(1) = 1 - 0.19687 - 0.34743 = 0.45570.$$

The first number is the auditor's chance of **detecting** the problem at all with this sample size; it is often uncomfortably far from 1 when p is small.

Problem 4

seed 20260925

For each situation decide whether Y is a binomial random variable in the sense of Definition 3.6. Where the sampling is from a very large population, treat "approximately binomial" (as in Example 3.6) as binomial.

(a) A mobile operator texts a retention offer to 25 subscribers chosen at random from its 3 million subscribers. Y = the number of the 25 who accept the offer.

- ○ Binomial
- ○ Not binomial: the number of trials is not fixed in advance
- ○ Not binomial: the trials are not independent, or p is not the same on every trial

◯ Not binomial: Y is not a count of successes in two-outcome trials

(b) Twelve exporters each report their percentage revenue growth for the year. Y = the average of the twelve growth rates.

◯ Binomial
◯ Not binomial: the number of trials is not fixed in advance
◯ Not binomial: the trials are not independent, or p is not the same on every trial
◯ Not binomial: Y is not a count of successes in two-outcome trials

(c) A food exporter in Guba ships 13 consignments by unrelated carriers on unrelated routes; each arrives late with probability 0.23, independently. Find the probability that at most one consignment is late, to at least four significant figures. _______________

Definition 3.6 has five conditions: a fixed number n of identical trials; two outcomes per trial; the same success probability p on every trial; independent trials; and Y = the number of successes. One failure is enough to rule the binomial model out.

(a) Answer: Binomial. Reason: fixed n = 25, two outcomes; sampling without replacement from 3 million subscribers changes the acceptance rate among those left by a negligible amount, so the trials are independent to an excellent approximation (Example 3.6).

(b) Answer: Not binomial: Y is not a count of successes in two-outcome trials. Reason: Y is an average of percentages, not a count of successes.

(c) This one satisfies all five conditions with $n = 13$, $p = 0.23$, $q = 0.77$. “At most one” is $\{Y = 0\} \cup \{Y = 1\}$, so

$$P(Y \leq 1) = q^{13} + 13\,p\,q^{12} = 0.03345 + 0.12989 = 0.16333.$$

Problem 5

seed 20260925

A microfinance lender receives 12 small-business loan applications this week from unrelated applicants. Each application arrives with the full set of documents with probability 0.87, independently of the others. Let Y be the number of complete applications.

Give every probability to at least four significant figures (a calculator or R’s `pbinom` is fine).

(a) $P(Y \leq 9) =$ _______________

(b) $P(Y > 9) =$ _______________

(c) $P(Y < 9) =$ _______________

Y is binomial with $n = 12$, $p = 0.87$, $q = 0.13$, so $p(y) = \binom{12}{y}\, 0.87^y \cdot 0.13^{12-y}$ (Definition 3.7).

(a) The cumulative probability is the sum of the probability function over $y = 0, 1, \ldots, 9$:

$$P(Y \leq 9) = \sum_{y=0}^{9} \binom{12}{y}\, 0.87^y \cdot 0.13^{12-y} = 0.19772,$$

which is pbinom(9, 12, 0.87) in R.

(b) $Y > 9$ is the complement of $Y \leq 9$, so $P(Y > 9) = 1 - 0.19772 = 0.80228$.

(c) For an integer-valued Y, $Y < 9$ is the same event as $Y \le 8$, so the strict inequality drops the term $p(9) = 0.13801$: $P(Y < 9) = P(Y \le 9) - p(9) = 0.19772 - 0.13801 = 0.05970$.

The difference between $\le$ and $<$ is a whole probability mass, not a rounding matter: for a discrete random variable, always ask which boundary value is included.

Problem 6

seed 20260925

A pension fund holds bonds issued by 10 unrelated companies in different industries. Over the next year each issuer defaults with probability 0.140, independently of the others. The fund's report tracks Y = the number of the 10 bonds that are **repaid in full** (do not default).

Give each answer to at least four significant figures.

(a) Y is binomial. What is its success probability p? ______________

(b) $P(Y = 8) =$ ______________

(c) $E(Y) =$ ______________

(d) $V(Y) =$ ______________

The book warns that the commonest error with the binomial model is failing to say which outcome is the success, and then using q where p belongs. Here Y counts repaid bonds, so a trial is a **success** when the bond is repaid, even though the number that catches the eye in the question is the default probability.

(a) $p = P(\text{repaid}) = 1 - 0.140 = 0.860$, and $q = 0.140$.

(b) $Y = 8$ means exactly 2 of the 10 bonds default. By Definition 3.7,

$$P(Y = 8) = \binom{10}{8} p^8 q^2 = 45 \times 0.860^8 \times 0.140^2 = 0.26391.$$

(The same number comes from the count of defaults, $n - Y$, which is binomial with success probability 0.140: $\binom{10}{2}$ equals $\binom{10}{8}$.)

(c) Theorem 3.7: $E(Y) = np = 10 \times 0.860 = 8.6$.

(d) $V(Y) = npq = 10 \times 0.860 \times 0.140 = 1.20400$, the same as the variance of the number of defaults, since npq is symmetric in p and q.

Problem 7

seed 20260925

A bank's mobile app lets new customers open an account remotely. Of the 16 people who started onboarding this morning, each finishes the process within 24 hours with probability 0.82, independently of the others. Let Y be the number who finish within 24 hours.

Give every probability to at least four significant figures.

(a) Exactly 13 finish. ______________

(b) At least 10 finish. ______________

(c) At least 9 but not more than 13 finish. ________________

(d) Fewer than 9 finish. ________________

Y is binomial with $n = 16$, $p = 0.82$, $q = 0.18$. Write $F(a) = P(Y \le a) = \sum_{y=0}^{a} p(y)$. Every event below is turned into one or two values of F by deciding which boundary values it contains.

(a) One term of Definition 3.7: $p(13) = \binom{16}{13} 0.82^{13} \cdot 0.18^3 = 560 \times 0.82^{13} \cdot 0.18^3 = 0.24751$.

(b) "At least 10" is $Y \ge 10$, whose complement is $Y \le 9$: $P(Y \ge 10) = 1 - F(9) = 1 - 0.01530 = 0.98470$.

(c) $9 \le Y \le 13$ contains both end points, so remove everything at or below 8 from everything at or below 13: $P = F(13) - F(8) = 0.56984 - 0.00356 = 0.56628$.

(d) "Fewer than 9" excludes 9 itself: $P(Y < 9) = F(8) = 0.00356$.

In R these are dbinom(13,16,0.82), 1-pbinom(9,16,0.82), pbinom(13,16,0.82)-pbinom(8,16,0.82) and pbinom(8,16,0.82).

Problem 8

seed 20260925

An energy company has raised enough capital to drill 7 exploration wells in separate prospects on the Caspian shelf. Each well is commercially successful with probability 0.28, independently of the others. Before the first well the company pays a fixed 22 million AZN to mobilise the drilling rig. Each successful well then costs 43 million AZN (drilling plus completion) and each dry well 17 million AZN. Let Y be the number of successful wells and C the total cost, in million AZN, of the programme.

Give each answer to at least four significant figures.

(a) $E(Y) =$ ________________

(b) $E(C) =$ ________________ million AZN

(c) $V(C) =$ ________________ (million AZN)^2

(d) The standard deviation of C is ________________ million AZN

Y is binomial with $n = 7$ and $p = 0.28$; the number of dry wells is $7 - Y$.

(a) Theorem 3.7: $E(Y) = np = 7 \times 0.28 = 1.96$, and $V(Y) = npq = 7 \times 0.28 \times 0.72 = 1.4112$.

(b) Write the cost as a linear function of Y:

$$C = 22 + 43Y + 17(7 - Y) = (22 + 119) + (43 - 17)Y = 141 + 26Y.$$

By Theorems 3.4 and 3.5 (constants and constant multiples come out of expectations),

$$E(C) = 141 + 26E(Y) = 141 + 26 \times 1.96 = 191.9600.$$

(c) A constant added to a random variable does not change its spread, and a constant multiple scales the variance by its square (Exercise 3.33(b)):

$$V(C) = 26^2 V(Y) = 676 \times 1.4112 = 953.9712.$$

(d) $\sigma_C = \sqrt{953.9712} = 30.8864$ million AZN.

Only the **difference** in cost between a successful and a dry well, 26 million AZN, carries the risk. The fixed cost and the base cost of every well are certain, however large they are.

Problem 9

seed 20260925

A distributor sells 5 card-payment terminals to a supermarket chain, taken at random from a large warehouse stock in which a proportion 0.19 of the units are defective. The chain returns every defective unit for repair. If Y is the number of defective terminals among the 5, the distributor's repair cost, in tens of AZN, is

$$C = 3Y^2 + 5Y + 6.$$

(The quadratic term reflects overtime once several units arrive together.)

Give each answer to at least four significant figures.

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

(b) The expected repair cost $E(C) =$ ____________ (tens of AZN)

(c) A junior analyst "estimates" the expected cost by plugging $E(Y)$ into the cost formula. By how much does $E(C)$ exceed that plug-in value? ____________

The warehouse stock is large, so Y is (approximately) binomial with $n = 5$, $p = 0.19$, $q = 0.81$. By Theorem 3.7, $\mu = np = 0.9500$ and $\sigma^2 = npq = 0.7695$.

(a) Theorem 3.6 says $\sigma^2 = E(Y^2) - \mu^2$, so

$$E(Y^2) = \sigma^2 + \mu^2 = 0.7695 + 0.9500^2 = 1.6720.$$

(b) Expectation is linear (Theorems 3.4 and 3.5), so it passes through the sum but not through the square:

$$E(C) = 3E(Y^2) + 5E(Y) + 6 = 3(1.6720) + 5(0.9500) + 6 = 15.7660.$$

(c) The plug-in value is $3\mu^2 + 5\mu + 6 = 13.4575$. The difference is

$$E(C) - C(\mu) = 3\left(E(Y^2) - \mu^2\right) = 3\sigma^2 = 3 \times 0.7695 = 2.3085.$$

Because the cost is convex in Y, its expectation is always above the cost at the expected count, by exactly the coefficient of Y^2 times the variance. Budgeting at $C(\mu)$ systematically under-provisions.

Problem 10

seed 20260925

An importer sells a shipment of 10 Wi-Fi routers to an internet provider at 130 AZN each. To win the contract it offers a double-your-money-back guarantee: for every defective router the provider receives 2×130 AZN back. Each router is defective with probability 0.12, independently of the others. Let G be the importer's net gain, in AZN, from the shipment (sales revenue minus refunds).

Give each answer to at least four significant figures.

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

(b) The standard deviation of G is ______________ AZN

(c) What is the probability that the net gain is less than 910 AZN? ______________

Let Y be the number of defective routers. Y is binomial with $n = 10$, $p = 0.12$, $q = 0.88$, so by Theorem 3.7 $E(Y) = 1.2000$ and $V(Y) = 1.0560$. The gain is a linear function of Y:

$$G = 10 \times 130 - 260\,Y = 1300 - 260\,Y.$$

(a) $E(G) = 1300 - 260E(Y) = 1300 - 260 \times 1.2000 = 988.00$ AZN.

(b) $V(G) = 260^2 V(Y)$, so $\sigma_G = 260\sqrt{1.0560} = 267.181$ AZN.

(c) Translate the event about G into an event about Y:

$$G < 910 \iff 1300 - 260Y < 910 \iff Y > \frac{1300 - 910}{260} = 1.5 \iff Y \geq 2.$$

Then by the complement,

$$P(Y \geq 2) = 1 - (p(0) + p(1)) = 1 - (0.27850 + 0.37977) = 0.34172.$$

The gain can only take the values $1300 - 260y$, so any threshold between two of them is the same event as a statement about a whole number of defectives.

Problem 11

seed 20260925

A bank's compliance officer audits a portfolio of 35 business accounts, of which 14 are flagged for incomplete customer-identification (KYC) documents. She selects 10 accounts at random **without replacement** and records Y = the number of flagged accounts in her sample.

Give every probability to at least four significant figures.

(a) The probability that the first account drawn is flagged. ______________

(b) The probability that the second account drawn is flagged, given that the first one was flagged. ______________

(c) Is the binomial distribution with $n = 10$ and p equal to your answer to (a) an appropriate model for Y?

○ Yes: the sample is a small fraction of the portfolio, so the draws are nearly independent and Y is approximately binomial

○ No: the sample is a large fraction of the portfolio, so the chance of a flagged account changes noticeably from draw to draw

○ Yes: sampling without replacement always produces a binomial count

○ No: each sampled account has more than two possible outcomes

(d) Using that binomial model regardless, compute $P(Y = 0)$. ______________

(e) Compute the exact $P(Y = 0)$ for sampling without replacement, using the multiplicative law draw by draw.

____

(a) Every account is equally likely to be first, so $P(S_1) = 14/35 = 0.40000$.

(b) After a flagged account is removed, 13 of the remaining 34 accounts are flagged: $P(S_2 \mid S_1) = 13/34 = 0.38235$. (As in Exercise 3.35, the **unconditional** $P(S_2)$ is still 0.40000; it is the conditional probability that moves.)

(c) Condition 4 of Definition 3.6 asks for independent trials, and independence would require (b) to equal (a). The two differ by 0.01765. Here the sample is 28.57% of the portfolio. Drawing one flagged account moves the chance for the next draw from 0.40000 to 0.38235, a material change, so the trials are not independent and the binomial model is not appropriate; the hypergeometric distribution of Section 3.7 is the right model.

So the answer is: No: the sample is a large fraction of the portfolio, so the chance of a flagged account changes noticeably from draw to draw.

(d) $P(Y = 0) = (1 - 0.40000)^{10} = 0.00605$.

(e) By the multiplicative law (Theorem 2.5) applied draw by draw,

$$P(Y = 0) = \frac{21}{35} \cdot \frac{21 - 1}{35 - 1} \cdots \frac{12}{26} = 0.00192.$$

Compare (d) with (e): when the sample is a tiny fraction of the portfolio they agree to several decimal places, which is exactly what "approximately binomial" means in Example 3.6; when it is a large fraction they do not.

Problem 12

seed 20260925

A leasing company has 18 equipment leases with unrelated small firms. Each lease is paid to term with probability 0.95, independently of the others. Let Y be the number of leases that default and $Y^* = 18 - Y$ the number paid to term.

Printed binomial tables often stop at $p = 0.5$ or list only a few values of p, so the risk analyst wants every answer expressed through Y, whose success probability is small.

Give each answer to at least four significant figures.

(a) $P(Y^* \geq 16)$ ______________

(b) $P(Y^* = 18)$ ______________

(c) $V(Y^*)$ ______________

(d) $E(Y)$ ______________

Y^* is binomial with $n = 18$ and success ("paid to term") probability 0.95. Exercise 3.54 shows that $Y = 18 - Y^*$ is then binomial with the same n and success probability $q = 0.05$: relabelling which outcome is called a success just swaps p and q in Definition 3.7, because $\binom{n}{n-y} = \binom{n}{y}$.

(a) $Y^* \geq 16 \iff 18 - Y \geq 16 \iff Y \leq 2$. With $p(y) = \binom{18}{y} 0.05^y \cdot 0.95^{18-y}$,

$$P(Y \leq 2) = p(0) + p(1) + p(2) = 0.39721 + 0.37631 + 0.16835 = 0.94187.$$

(b) $Y^* = 18 \iff Y = 0$: $P = 0.95^{18} = 0.39721$.

(c) $V(Y^*) = npq = 18 \times 0.95 \times 0.05 = 0.8550$. Since $V(18 - Y) = V(Y)$, the number of defaults has the same variance.

(d) $E(Y) = nq = 18 \times 0.05 = 0.9000$, and $E(Y^*) = 18 - E(Y)$.

Problem 13

seed 20260925

A fintech company signs up shops for its QR-code payment service. A sales agent visits shops one at a time; each shop visited signs up with probability 0.12, independently of the others. The agent is paid a commission of 35 AZN per shop signed. The regional manager wants to plan a day of visits so that the probability of **at least one** sign-up is at least 0.98.

(a) What is the smallest number of shops the agent must visit? ________________

(b) With that number of visits, what is the probability of at least one sign-up? (at least four significant figures)

____

(c) With that number of visits, what is the agent's expected commission for the day, in AZN? ________________

With n visits, the number of sign-ups Y is binomial with parameters n and $p = 0.12$. By the complement, $P(Y \geq 1) = 1 - p(0) = 1 - q^n = 1 - 0.88^n$.

(a) We need $1 - 0.88^n \geq 0.98$, i.e. $0.88^n \leq 0.02$. Taking logarithms (and remembering that $\ln 0.88 < 0$, so the inequality turns round when we divide by it):

$$n \geq \frac{\ln 0.02}{\ln 0.88} = \frac{-3.91202}{-0.12783} = 30.603.$$

The smallest whole number satisfying this is $n = 31$. Check: with $n = 30$, $1 - 0.88^{30} = 0.97840 < 0.98$, which is not enough.

(b) $1 - 0.88^{31} = 0.98099$.

(c) The commission is $35Y$, so by Theorem 3.7, $E(35Y) = 35 \cdot np = 35 \times 31 \times 0.12 = 35 \times 3.7200 = 130.20$ AZN.

Problem 14

seed 20260925

A pension fund is reviewing an equity manager who beat her benchmark (after fees) in 8 of the last 14 months. The fund's analysts calculate that a manager with **no skill** would beat this benchmark in any given month with probability 0.47, independently from month to month. Let Y be the number of months out of 14 in which a no-skill manager beats the benchmark.

(a) The expected number of months in which a no-skill manager beats the benchmark. ________________

(b) The probability that a no-skill manager beats the benchmark in **at least** 8 of the 14 months (at least four significant figures). ________________

(c) Using 0.05 as the dividing line, is her record unusual for a manager with no skill?

○ Yes: under the no-skill model a record this good has probability below 0.05, so the no-skill explanation is hard to believe

○ No: under the no-skill model a record at least this good has probability 0.05 or more, so it is not unusual for a manager with no skill

Under the no-skill model, Y is binomial with $n = 14$ and $p = 0.47$, $q = 0.53$.

(a) Theorem 3.7: $E(Y) = np = 14 \times 0.47 = 6.58$. Her 8 months is to be compared with this.

(b) The relevant probability is not $P(Y = 8)$ (which is only 0.15849 and would be small for any single value) but the probability of a record **at least as good** as hers, exactly as in Example 3.8:

$$P(Y \geq 8) = 1 - P(Y \leq 7) = 1 - 0.68947 = 0.31053.$$

(c) Since 0.31053 is not below 0.05, a manager with no skill would produce a record at least this good fairly often. The record is not strong evidence of skill. Answer: No: under the no-skill model a record at least this good has probability 0.05 or more, so it is not unusual for a manager with no skill.

Problem 15

seed 20260925

A bank is considering a mobile savings "pot" product. It surveys 18 customers chosen at random from its large retail base, and 12 say they would use it. Let p be the true proportion of the bank's customers who would use the product, so that the number in the sample who say yes is binomial with $n = 18$ and unknown p.

Give each answer to at least four significant figures.

(a) The value of p that maximises the probability of observing exactly 12 yes answers. ________________

(b) That maximised probability, $P(Y = 12)$ evaluated at your answer to (a). ________________

(c) The product manager claims $p = 0.35$. Evaluate $P(Y = 12)$ at $p = 0.35$. ________________

(d) How many times more probable is the observed result under the estimate from (a) than under the manager's claim? (the ratio of (b) to (c)) ________________

(a) As in Example 3.10, $P(Y = 12) = \binom{18}{12} p^{12}(1-p)^6$. The binomial coefficient does not depend on p, and $\ln$ is increasing, so maximise

$$\ell(p) = 12 \ln p + 6 \ln(1-p), \qquad \ell'(p) = \frac{12}{p} - \frac{6}{1-p} = 0 \implies p = \frac{12}{18} = 0.66667.$$

$\ell''(p) = -12/p^2 - 6/(1-p)^2 < 0$, so this is a maximum. The estimate is the sample proportion of yes answers.

(b) $\binom{18}{12} = 18564$, so $P(Y = 12) = 18564 \times 0.66667^{12} \times 0.33333^6 = 0.19627$.

(c) $18564 \times 0.35^{12} \times 0.65^6 = 0.0047312$.

(d) $0.19627/0.0047312 = 41.484$. The binomial coefficient cancels in the ratio. By construction the ratio is at least 1; the further it is above 1, the less well the manager's figure explains what the survey actually saw. (Chapter 9 turns this comparison into a formal method.)

Problem 16

seed 20260925

A boutique hotel in Sheki has 14 rooms. For a festival weekend it accepts 18 non-refundable one-room bookings, 4 more than it has rooms, because some guests never arrive. Each booked guest arrives with probability 0.84, independently of the others (the bookings come from unrelated travellers). Let Y be the number of guests who arrive.

Give each answer to at least four significant figures.

(a) The probability that more guests arrive than there are rooms. ________________

(b) $E(Y) =$ ______________

(c) The expected number of arriving guests who cannot be given a room. ______________

(d) The expected number of rooms left empty. ______________

Y is binomial with $n = 18$, $p = 0.84$, $q = 0.16$.

(a) Over-capacity is $Y > 14$, i.e. $Y \geq 15$:

$$P(Y > 14) = p(15) + \cdots + p(18) = 0.24448 + 0.24066 + 0.14864 + 0.04335 = 0.67713.$$

(b) Theorem 3.7: $E(Y) = np = 18 \times 0.84 = 15.1200$.

(c) The number turned away is $g(Y) = \max(Y - 14, 0)$, which is $y - 14$ for $y > 14$ and 0 otherwise. By Theorem 3.2, $E[g(Y)] = \sum_y g(y)p(y)$:

$$E[\max(Y - 14, 0)] = 1 \cdot 0.24448 + 2 \cdot 0.24066 + 3 \cdot 0.14864 + 4 \cdot 0.04335 = 1.34513.$$

(d) The rooms left empty are $\max(14 - Y, 0)$. For every y, $\max(14 - y, 0) = 14 - y + \max(y - 14, 0)$ (check the cases $y \leq 14$ and $y > 14$). Taking expectations term by term,

$$E[\max(14 - Y, 0)] = 14 - E(Y) + E[\max(Y - 14, 0)] = 14 - 15.1200 + 1.34513 = 0.22513.$$

Overbooking trades (c) against (d): each extra booking lowers the expected number of empty rooms and raises the expected number of guests who must be compensated and sent elsewhere.

Problem 17

seed 20260925

A bank is considering buying a pool of several thousand consumer loans from a microfinance originator. Before signing, its credit-risk team audits a random sample of loan files for documentation defects and buys the pool only if the number of defective files in the sample is small. Two audit plans are on the table:

- **Plan 1:** audit 11 files; buy if at most 1 is defective.
- **Plan 2:** audit 24 files; buy if at most 3 are defective.

The bank calls the pool **good** if 0.03 of its files are defective and **bad** if 0.19 are. The pool is so large that the number of defective files in the sample may be treated as binomial.

Give every probability to at least four significant figures.

(a) The probability that the bank buys a **good** pool under Plan 1 ______________ and under Plan 2 ______________.

(b) The probability that the bank buys a **bad** pool under Plan 1 ______________ and under Plan 2 ______________.

(c) Which plan gives the bank the smaller probability of buying a bad pool?

○ Plan 1
○ Plan 2

(d) Which plan gives the smaller probability of wrongly rejecting a good pool (the risk borne by an honest originator)?

○ Plan 1
○ Plan 2

Under a plan that audits n files and buys when at most c are defective, the number of defective files Y is binomial with parameters n and the pool's defect rate p, and

$$P(\text{buy}) = P(Y \le c) = \sum_{y=0}^{c} \binom{n}{y} p^y (1-p)^{n-y}.$$

(a) Good pool, $p = 0.03$:

- Plan 1: $P(Y \le 1) = p(0) + p(1) = 0.71530 + 0.24335 = 0.95865$.
- Plan 2: $P(Y \le 3) = \sum_{y=0}^{3} \binom{24}{y} 0.03^y (1-0.03)^{24-y} = 0.99468$.

(b) Bad pool, $p = 0.19$:

- Plan 1: $P(Y \le 1) = p(0) + p(1) = 0.09848 + 0.25410 = 0.35257$.
- Plan 2: $P(Y \le 3) = \sum_{y=0}^{3} \binom{24}{y} 0.19^y (1-0.19)^{24-y} = 0.30502$.

(c) Compare 0.35257 (Plan 1) with 0.30502 (Plan 2): the smaller is Plan 2.

(d) The probabilities of rejecting a good pool are $1 - 0.95865 = 0.04135$ (Plan 1) and $1 - 0.99468 = 0.00532$ (Plan 2): the smaller is Plan 2.

A larger sample does not automatically protect the buyer: raising the allowed number of defects c pushes both acceptance probabilities up. The two risks pull against each other, and a plan is chosen by deciding how much of each the bank and the originator will tolerate, which is a binomial calculation at two values of p.

Problem 18

seed 20260925

A regional airline flies a 19-seat turboprop between Baku and Lankaran. A ticket costs 115 AZN and is non-refundable, so the airline keeps the fare whether or not the passenger shows up. Each ticket holder shows up with probability 0.91, independently of the others. If more ticket holders show up than there are seats, each passenger who cannot board is paid 250 AZN compensation. The airline will sell between 19 and 24 tickets and wants to maximise its expected revenue net of compensation.

Let Y_n be the number of ticket holders who show up when n tickets are sold. Give each answer to at least four significant figures.

(a) If 21 tickets are sold, the expected number of passengers who cannot board. ____________

(b) If 21 tickets are sold, the expected net revenue, in AZN. ____________

(c) The number of tickets, between 19 and 24, that maximises expected net revenue. ____________

(d) That maximum expected net revenue, in AZN. ____________

With n tickets sold, Y_n is binomial with parameters n and $p = 0.91$. The number who cannot board is $\max(Y_n - 19, 0)$, and the net revenue is

$$R_n = 115\,n - 250 \max(Y_n - 19, 0), \qquad E(R_n) = 115\,n - 250\,E[\max(Y_n - 19, 0)],$$

using Theorem 3.2 for the expectation of a function of Y_n and linearity.

(a) With $n = 21$ only $y = 20$ (one passenger bumped) and $y = 21$ (two bumped) contribute:

$$E[\max(Y_{21} - 19, 0)] = 1 \cdot p(20) + 2 \cdot p(21) = 0.28661 + 2 \times 0.13800 = 0.5626.$$

(b) $E(R_{21}) = 2415 - 250 \times 0.5626 = 2274.35$ AZN.

(c) and **(d)** Repeat for every n from 19 to 24:

- $n = 19$: expected bumped 0.0000, expected net revenue 2185.00
- $n = 20$: expected bumped 0.1516, expected net revenue 2262.09
- $n = 21$: expected bumped 0.5626, expected net revenue 2274.35
- $n = 22$: expected bumped 1.2069, expected net revenue 2228.26
- $n = 23$: expected bumped 1.9983, expected net revenue 2145.42
- $n = 24$: expected bumped 2.8625, expected net revenue 2044.38

The largest is at $n = 21$, with expected net revenue 2274.35 AZN.

A quicker way to read the table: selling one more ticket brings in 115 AZN for certain, and it causes one extra bumped passenger exactly when the new holder shows up and the other n holders have already filled the plane, which has probability $p \cdot P(Y_n \geq 19)$. So the extra ticket pays while $115 > 250 \cdot 0.91 \cdot P(Y_n \geq 19)$, and the optimum is where that first fails (or the upper limit of 24).

Problem 19

seed 20260925

A bank packages 22 loans of 1 million AZN each, made to unrelated firms in different sectors and regions, into a pool and sells it to investors in three slices (tranches). Over the coming year each loan defaults with probability 0.095, independently of the others, and each default costs the pool 0.60 million AZN. Losses are absorbed in order of seniority:

- the **junior** tranche absorbs the losses from the first 2 defaults;
- the **middle** tranche absorbs the losses from the next 4 defaults (defaults number 3 to 6);
- the **senior** tranche absorbs any losses beyond 6 defaults.

Let Y be the number of defaults. Give every answer to at least four significant figures.

(a) The probability that the pool suffers no default. ________________

(b) The probability that the junior tranche is wiped out completely. ________________

(c) The probability that the senior tranche loses anything. ________________

(d) The expected loss of the middle tranche, in million AZN. ________________

Y is binomial with $n = 22$, $p = 0.095$, $q = 0.905$.

(a) $P(Y = 0) = q^{22} = 0.905^{22} = 0.11124$.

(b) The junior tranche is gone once there are at least 2 defaults: $P(Y \geq 2) = 1 - P(Y \leq 1) = 1 - 0.36814 = 0.63186$.

(c) The senior tranche is touched only by a default beyond the 6th:
$P(Y > 6) = 1 - P(Y \leq 6) = 1 - 0.996712 = 0.003288$.

(d) The number of defaults falling on the middle tranche is

$$M = \min\left(\max(Y-2,0),\ 4\right),$$

which is 0 when $Y \le 2$, $Y-2$ when $2 < Y \le 6$, and 4 when $Y > 6$. Write M as a sum of indicators: the jth default of the middle tranche happens exactly when $Y \ge 2+j$, so $M = \sum_{j=1}^{4} I(Y \ge 2+j)$ and

$$E(M) = \sum_{j=1}^{4} P(Y \ge 2+j) = P(Y \ge 3) + \cdots + P(Y \ge 6) = 0.34870 + 0.15054 + 0.05173 + 0.01439 = 0.56536.$$

(The same number comes from $\sum_y m(y)p(y)$ by Theorem 3.2.) Each of those defaults costs 0.60 million AZN, so the expected loss of the middle tranche is $0.60 \times 0.56536 = 0.33922$ million AZN.

For comparison, the whole pool's expected loss is $0.60 \cdot np = 1.2540$ million AZN (Theorem 3.7), so the middle tranche carries about 27.1% of it. Every tranche probability here rests on the independence of the 22 defaults; with a common cause driving defaults, the chance of many defaults together, and hence (c), would be far larger.

Problem 20

seed 20260925

A bank's Ganja branch must fill 3 teller posts before a new service hall opens. It makes job offers to shortlisted candidates, each of whom accepts with probability 0.68, independently of the others, and it will take on every candidate who accepts. HR wants the number of offers to be the smallest that makes the probability of at least 3 acceptances at least 0.90. Each accepted offer costs 375 AZN in training and onboarding.

Give each probability to at least four significant figures.

(a) The number of offers HR should make. ________________

(b) With that many offers, the probability of at least 3 acceptances. ________________

(c) With one offer fewer, the probability of at least 3 acceptances. ________________

(d) With the number of offers in (a), the expected onboarding cost, in AZN. ________________

With n offers the number of acceptances Y is binomial with parameters n and $p = 0.68$. We need the smallest n with

$$P(Y \ge 3) = 1 - P(Y \le 2) = 1 - \sum_{y=0}^{2} \binom{n}{y} 0.68^y \cdot 0.32^{n-y} \ge 0.90.$$

Unlike "at least one", this has no closed-form solution for n, so try $n = 3, 3+1, \ldots$ in turn. $P(Y \ge 3)$ increases with n, so the first n that works is the answer.

(a) and **(c)** At $n = 5$: $P(Y \le 2) = 0.19053$, so $P(Y \ge 3) = 0.80947 < 0.90$, not enough.

(b) At $n = 6$: $P(Y \le 2) = 0.08749$, so $P(Y \ge 3) = 0.91251 \ge 0.90$. HR should make 6 offers.

(d) All acceptances are hired, so the cost is $375Y$ and, by Theorem 3.7,
$E(375Y) = 375 \cdot np = 375 \times 6 \times 0.68 = 375 \times 4.0800 = 1530.00$ AZN.

The expected number of acceptances, 4.0800, exceeds the 3 posts by 1.0800: making sure the posts are filled with high probability means expecting, on average, to hire more people than strictly needed.

Answer key

Problem	Answers
1	(a) 0.281512, (b) 0.021834
2	(a) 4.48, (b) 3.0464, (c) 1.74539, (d) 9.52
3	(a) 0.803126, (b) 0.4557
4	(a) Binomial, (b) Not binomial ... outcome trials, (c) 0.163334
5	(a) 0.197717, (b) 0.802283, (c) 0.0597017
6	(a) 0.86, (b) 0.26391, (c) 8.6, (d) 1.204
7	(a) 0.247506, (b) 0.984697, (c) 0.566281, (d) 0.00356318
8	(a) 1.96, (b) 191.96, (c) 953.971, (d) 30.8864
9	(a) 1.672, (b) 15.766, (c) 2.3085
10	(a) 988, (b) 267.181, (c) 0.341725
11	(a) 0.4, (b) 0.382353, (c) No: the ... draw to draw, (d) 0.00604662, (e) 0.00192133
12	(a) 0.941871, (b) 0.397214, (c) 0.855, (d) 0.9
13	(a) 31, (b) 0.980991, (c) 130.2
14	(a) 6.58, (b) 0.310526, (c) No: under ... with no skill
15	(a) 0.666667, (b) 0.196268, (c) 0.00473117, (d) 41.484
16	(a) 0.677131, (b) 15.12, (c) 1.34513, (d) 0.225133
17	(a) 0.958651, (b) 0.994679, (c) 0.352572, (d) 0.305016, (e) Plan 2, (f) Plan 2
18	(a) 0.562603, (b) 2274.35, (c) 21, (d) 2274.35
19	(a) 0.111242, (b) 0.631857, (c) 0.00328785, (d) 0.339216
20	(a) 6, (b) 0.912507, (c) 0.809474, (d) 1530