Problem Set — Week 3, Session 1 · Solutions

Mathematical Statistics I · Wackerly §2.7 · ADA University

> 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 20260923

A branch classifies the 282 loan applications it received last quarter by whether the loan was secured and by whether it was approved.

	Secured	Unsecured
Approved:	138	63
Refused:	38	43

One of the 282 applications is drawn at random, so that every application is equally likely. Let A be the event that it was approved and S the event that it was secured.

Give every probability to at least four significant figures.

(a) $P(A \cap S) =$ ____________

(b) $P(A \mid S) =$ ____________

(c) $P(S \mid A) =$ ____________

Solution

Every application is one equally likely sample point, so a probability is a count divided by 282.

(a) The cell "approved and secured" holds 138 applications, so $P(A \cap S) = 138/282 = 0.4894$.

(b) By Definition 2.9,

$$P(A \mid S) = \frac{P(A \cap S)}{P(S)} = \frac{138/282}{176/282} = \frac{138}{176} = 0.7841.$$

The division by 282 cancels: conditioning on S simply throws away every application that was not secured and re-counts within the 176 that remain.

(c) The same definition with the roles exchanged: $P(S \mid A) = 138/201 = 0.6866$.

Parts (b) and (c) have the same numerator and different denominators, which is why they are different numbers. "The loan was approved, given that it was secured" and "the loan was secured, given that it was approved" are different questions, and confusing them is the most expensive mistake in this chapter.

Problem 2

seed 20260923

Let A be the event that a mid-sized exporter raises new bank debt this year and B the event that it opens a new foreign subsidiary this year. For the firms on a regulator's register,

$$P(A) = 0.69, \qquad P(B) = 0.50, \qquad P(A \cap B) = 0.4450.$$

Give every probability to at least four significant figures.

(a) $P(A \mid B) =$ ____________

(b) $P(B \mid A) =$ ____________

(c) $P(A \cup B) =$ ____________

(d) Are A and B independent, dependent, or mutually exclusive?

- ○ Dependent
- ○ Independent
- ○ Mutually exclusive

Solution

(a) Definition 2.9: $P(A \mid B) = P(A \cap B)/P(B) = 0.4450/0.50 = 0.8900$.

(b) $P(B \mid A) = 0.4450/0.69 = 0.6449$.

(c) The additive law: $P(A \cup B) = 0.69 + 0.50 - 0.4450 = 0.7450$. Subtracting the intersection once repairs the double count of the firms in both events.

(d) Definition 2.10 offers three equivalent tests, and any one of them settles it. Here

$$P(A)P(B) = 0.69 \times 0.50 = 0.3450 \neq 0.4450 = P(A \cap B),$$

so the events are **dependent**. Equivalently $P(A \mid B) = 0.8900$ while $P(A) = 0.69$: learning that the firm opened a subsidiary changes what we would say about its borrowing.

They are certainly not mutually exclusive — that would require $P(A \cap B) = 0$, and here the two events happen together for a good share of the register. Mutually exclusive events with positive probabilities are in fact as far from independent as events can be, because then $P(A \mid B) = 0 \neq P(A)$.

Problem 3

seed 20260923

An insurer classifies its 270 policies in force by line of business and by whether a claim was made during the year.

	Claim	No claim	Total
Motor:	88	70	158

Household: 39 73 112

Total: 127 143 270

One policy is drawn at random. Let C be the event that a claim was made and M the event that the policy is a motor policy. Give every answer to at least four significant figures.

(a) $P(C) =$ ______________

(b) $P(C \mid M) =$ ______________

(c) C and M are not independent. How many motor policies would have had to make a claim — that is, what would the top-left cell have to be — for C and M to be independent? (A non-integer answer is expected; do not round it.) ______________

Solution

(a) $P(C) = 127/270 = 0.4704$.

(b) Conditioning on M restricts attention to the 158 motor policies, of which 88 claimed:
$P(C \mid M) = 88/158 = 0.5570$.

Because $0.5570 \neq 0.4704$, the first form of Definition 2.10 fails and the two classifications are dependent. The same conclusion follows from the third form: $P(M)P(C) = 0.5852 \times 0.4704 = 0.2753$, while $P(M \cap C) = 0.3259$.

(c) Write x for the top-left cell. Independence requires

$$\frac{x}{270} = P(M \cap C) = P(M)P(C) = \frac{158}{270} \cdot \frac{127}{270},$$

so $x = 158 \times 127/270 = 74.32$.

That expression — row total times column total divided by the grand total — is the **expected cell count** under independence. It is worth remembering: the whole of the chi-square test of independence, which this course reaches in the second semester, is built on comparing each observed cell with exactly this number.

Problem 4

seed 20260923

A fund holds 68 corporate bonds, of which 42 are investment grade and 26 are high yield. Two of them are selected at random for an audit, one after the other and **without replacement**. Let G_1 be the event that the first bond drawn is investment grade and G_2 the event that the second one is.

Give every probability to at least four significant figures.

(a) $P(G_1) =$ ______________

(b) $P(G_2 \mid G_1) =$ ______________

(c) $P(G_1 \cap G_2) =$ ______________

(d) $P(G_2 \mid \bar{G}_1) =$ ______________

Solution

(a) All 68 bonds are equally likely to be drawn first, so $P(G_1) = 42/68 = 0.6176$.

(b) Once an investment-grade bond has been removed, 67 bonds remain and 41 of them are investment grade: $P(G_2 \mid G_1) = 41/67 = 0.6119$.

(c) Definition 2.9, written as a product:

$$P(G_1 \cap G_2) = P(G_1)\, P(G_2 \mid G_1) = \frac{42}{68} \cdot \frac{41}{67} = 0.3780.$$

(d) If the first bond was high yield, all 42 investment-grade bonds are still there among the 67 that remain: $P(G_2 \mid \bar{G}_1) = 42/67 = 0.6269$.

Since (b) and (d) differ, G_1 and G_2 are dependent — which is exactly what sampling without replacement means. Worth noting, though, that the **unconditional** probability of G_2 is 0.6176, the same as $P(G_1)$: before the first bond is seen, the second draw is just as likely to be investment grade as the first. Dependence is about what one draw tells you about the other, not about the two draws having different unconditional behaviour.

Problem 5

seed 20260923

In a microfinance portfolio, B is the event that a borrower is taking a loan for the first time and A is the event that the borrower falls into arrears during the first year. The lender's records give

$$P(B) = 0.55, \qquad P(A \mid B) = 0.42, \qquad P(A) = 0.3600.$$

Give every probability to at least four significant figures.

(a) $P(A \cap B) =$ ______________

(b) $P(A \cup B) =$ ______________

(c) $P(B \mid A) =$ ______________

Solution

(a) Definition 2.9 rearranged is the multiplicative law: $P(A \cap B) = P(B)\, P(A \mid B) = 0.55 \times 0.42 = 0.2310$.

This is the useful direction in practice. A conditional probability is usually what the data hand you — the arrears rate **among first-time borrowers** is something the lender measures directly — and the intersection is what you want.

(b) $P(A \cup B) = 0.3600 + 0.55 - 0.2310 = 0.6790$.

(c) Now condition the other way round: $P(B \mid A) = P(A \cap B)/P(A) = 0.2310/0.3600 = 0.6417$.

So 0.6417 of the borrowers in arrears are first-time borrowers, while 0.42 of first-time borrowers fall into arrears. The two figures answer different questions and a supervisory report that swaps them says something false about the portfolio.

Incidentally, the borrowers in arrears who are **not** first-time borrowers have probability $0.3600 - 0.2310 = 0.1290$.

Problem 6

seed 20260923

Two trading desks close their books independently. Let A be the event that the Baku desk reports a profit on a given day and B the event that the Istanbul desk does, with

$$P(A) = 0.74, \qquad P(B) = 0.50,$$

and A and B independent. Give every probability to at least four significant figures.

(a) $P(A \cap B) =$ ____________

(b) $P(\bar{A} \cap B) =$ ____________

(c) $P(\bar{A} \cap \bar{B}) =$ ____________

(d) $P(A \mid \bar{B}) =$ ____________

Solution

(a) Independence gives the third form of Definition 2.10 directly: $P(A \cap B) = 0.74 \times 0.50 = 0.3700$.

(b) Split B along A: $B = (A \cap B) \cup (\bar{A} \cap B)$, a union of mutually exclusive events, so $P(\bar{A} \cap B) = P(B) - P(A \cap B) = 0.50 - 0.3700 = 0.1300$. Notice this equals $P(\bar{A})P(B) = 0.2600 \times 0.50$, so $\bar{A}$ and B are independent too.

(c) By the same argument applied to $\bar{B}$,

$$P(\bar{A} \cap \bar{B}) = P(\bar{A})\, P(\bar{B}) = 0.2600 \times 0.5000 = 0.1300.$$

A second route: $\bar{A} \cap \bar{B} = \overline{A \cup B}$ by De Morgan's law, and $P(A \cup B) = 0.8700$, so the answer is $1 - 0.8700 = 0.1300$. The two routes agreeing is not a coincidence — it is the algebraic content of the claim that independence passes to complements.

(d) Independence of A and $\bar{B}$ means conditioning on $\bar{B}$ tells us nothing: $P(A \mid \bar{B}) = P(A) = 0.74$.

The general statement behind all of this: if A and B are independent, then so are A and $\bar{B}$, $\bar{A}$ and B, and $\bar{A}$ and $\bar{B}$. Independence is a property of the **pair of partitions**, not of the particular events named.

Problem 7

seed 20260923

A rating agency will put a bank on review with one of exactly three outcomes at the next committee: an upgrade (A), a downgrade (B), or no change (C). Exactly one of the three occurs, and

$$P(A) = 0.44, \qquad P(B) = 0.30.$$

Give every probability to at least four significant figures.

(a) $P(C) =$ ____________

(b) $P(A \cup B) =$ ____________

(c) $P(A \mid B) =$ ____________

(d) Are A and B independent?

- ○ Dependent
- ○ Independent
- ○ Not enough information

(e) $P(A \mid \bar{C}) =$ ____________

Solution

(a) The three outcomes are mutually exclusive and exhaust the sample space, so their probabilities sum to 1 and $P(C) = 1 - 0.44 - 0.30 = 0.2600$.

(b) Mutually exclusive events have $P(A \cap B) = 0$, so the additive law loses its last term: $P(A \cup B) = 0.44 + 0.30 = 0.7400$.

(c) $P(A \mid B) = P(A \cap B)/P(B) = 0/0.30 = 0$. Once the bank has been downgraded it cannot also have been upgraded.

(d) Dependent, and strongly so. Independence would require $P(A \mid B) = P(A) = 0.44$, but the conditional probability is 0. In the product form, $P(A)P(B) = 0.1320 \neq 0 = P(A \cap B)$.

This is worth stating in general, because the two words are often confused: two events with positive probability that are **mutually exclusive** are never independent. Mutual exclusivity says that one occurring rules the other out — about as informative as an event can be — whereas independence says that one occurring tells you nothing about the other.

(e) $\bar{C} = A \cup B$, and $A \cap \bar{C} = A$, so

$$P(A \mid \bar{C}) = \frac{P(A)}{P(A \cup B)} = \frac{0.44}{0.7400} = 0.5946.$$

Conditioning on "the rating changed" rescales the two remaining outcomes so that they again sum to one.

Problem 8

seed 20260923

A data centre reaches the internet over two links that fail independently of one another. On a given day the primary link is available with probability 0.965 and the backup link with probability 0.855.

Let A_1 and A_2 be the events that the primary and the backup link are available. Give every probability to at least four significant figures.

(a) $P(A_1 \cap A_2) =$ ______________

(b) $P(\bar{A}_1 \cap \bar{A}_2) =$ ______________

(c) $P(\text{at least one link is available}) =$ ______________

(d) $P(\text{exactly one link is available}) =$ ______________

Solution

(a) Independence: $P(A_1 \cap A_2) = 0.965 \times 0.855 = 0.8251$.

(b) Independence passes to complements, so $P(\bar{A}_1 \cap \bar{A}_2) = 0.0350 \times 0.1450 = 0.0051$.

(c) “At least one” is the complement of “neither”, and Theorem 2.7 turns one into the other: $1 - 0.0051 = 0.9949$.

Going through the complement is the standard move and it is worth acquiring as a reflex. The direct route needs the additive law, $P(A_1) + P(A_2) - P(A_1 \cap A_2) = 0.965 + 0.855 - 0.8251$, which gives the same number by a longer road; with ten links instead of two the direct route would have 1023 terms and the complement would still have one.

(d) Exactly one means the primary works and the backup does not, or the other way round. These two events are mutually exclusive, so the probabilities add:

$$P = 0.965 \times 0.1450 + 0.0350 \times 0.855 = 0.1698.$$

Equivalently, “at least one” minus “both”: $0.9949 - 0.8251 = 0.1698$.

The design point: the second link raises availability from 0.965 to 0.9949, and it does so **only because the failures are independent**. Two links drawing power from the same substation fail together, and the calculation above then overstates the redundancy badly.

Problem 9

seed 20260923

For borrowers in a distressed portfolio, let B be the event that a borrower misses a scheduled payment during the year and A the event that the borrower defaults outright. Every default begins with a missed payment, so $A \subset B$. The portfolio has

$$P(A) = 0.2368, \qquad P(B) = 0.74.$$

Give every probability to at least four significant figures.

(a) $P(A \cap B) =$ ______________

(b) $P(A \mid B) =$ ______________

(c) $P(B \mid A) =$ ______________

(d) $P(A \cup B) =$ ______________

(e) $P(\bar{A} \cap B) =$ ______________

Solution

Everything here follows from $A \subset B$, which makes $A \cap B = A$.

(a) $P(A \cap B) = P(A) = 0.2368$.

(b) $P(A \mid B) = P(A)/P(B) = 0.2368/0.74 = 0.3200$. Among borrowers who miss a payment, this fraction go on to default — the quantity a workout desk actually plans around.

(c) $P(B \mid A) = P(A)/P(A) = 1$. A default is certain to have involved a missed payment; that is what containment says.

(d) $A \cup B = B$, so the union has probability 0.74. You can also run the additive law and watch the terms cancel: $0.2368 + 0.74 - 0.2368 = 0.74$.

(e) The borrowers who miss a payment but recover: $P(B) - P(A) = 0.74 - 0.2368 = 0.5032$.

Note that A and B are dependent unless $P(B) = 1$: a nested pair is about as far from independent as a pair can be, because B occurring is necessary for A.

Problem 10

seed 20260923

Two analysts call the direction of the manat–dollar rate for the coming quarter. They work from different data and their calls are independent. Analyst A is right with probability 0.87 and analyst B with probability 0.68. Write A and B for the events that each is right.

Give every probability to at least four significant figures.

(a) $P(A \cap B) =$ ______________

(b) $P(\text{at least one of them is right}) =$ ______________

(c) $P(A \mid \text{at least one is right}) =$ ______________

(d) $P(\text{both are right} \mid \text{at least one is right}) =$ ______________

Solution

(a) $P(A \cap B) = 0.87 \times 0.68 = 0.5916$.

(b) Neither is right with probability $0.1300 \times 0.3200 = 0.0416$, so at least one is right with probability $1 - 0.0416 = 0.9584$.

(c) Write $L = A \cup B$ for "at least one is right". Since $A \subset L$, we have $A \cap L = A$ and

$$P(A \mid L) = \frac{P(A)}{P(L)} = \frac{0.87}{0.9584} = 0.9078.$$

(d) Similarly $(A \cap B) \subset L$, so

$$P(A \cap B \mid L) = \frac{0.5916}{0.9584} = 0.6173.$$

Both conditional probabilities are larger than their unconditional counterparts, and that is the point of the exercise: "at least one was right" is evidence in favour of any particular one having been right. Conditioning is not a formality, it moves the numbers.

A caution for part (d): A and B are independent, but they are **not** independent given L. Conditioning on an event that both feed into creates dependence between them where there was none — knowing that at least one call was right, learning that A was wrong tells you B was right. Independence is not preserved by conditioning.

Problem 11

seed 20260923

A bank reviews every card account opened last year. Branch A opened 890 accounts, of which 71 were later flagged for fraud; branch B opened 570 accounts, of which 29 were flagged. One account is drawn at random from all 1460 of them. Let F be the event that it was flagged and A the event that it was opened at branch A.

Give every probability to at least four significant figures.

(a) $P(F) =$ ______________

(b) $P(F \mid A) =$ ______________

(c) $P(A \mid F) =$ ______________

Solution

Because every account is equally likely, each probability is a count over a count — which is all Definition 2.9 ever is once the sample points are equally likely.

(a) $71 + 29 = 100$ accounts were flagged out of 1460, so $P(F) = 0.0685$.

(b) Conditioning on A restricts attention to branch A's 890 accounts: $P(F \mid A) = 71/890 = 0.0798$.

(c) Conditioning on F restricts attention to the 100 flagged accounts, of which 71 came from branch A: $P(A \mid F) = 71/100 = 0.7100$.

The three numbers answer three different questions and the second and third are the pair most often confused in practice. $P(F \mid A) = 0.0798$ is the fraud rate at branch A — a statement about how branch A performs. $P(A \mid F) = 0.7100$ is the share of the bank's fraud cases that originated at branch A — a statement about where the caseload sits. A branch can have the higher rate and the smaller share, simply by being smaller; here branch A holds 0.6096 of all accounts, and that base rate is what carries $P(A \mid F)$ away from $P(F \mid A)$.

Section 2.10 will give Bayes' rule, which computes (c) from (b) and the branch sizes without ever writing down the counts. It is the same arithmetic; having done it once with counts makes the formula legible rather than magical.

Problem 12

seed 20260923

A reviewer with no real preference ranks 9 funds from first to last, assigning every one of the $9!$ possible orderings the same probability. Two of the funds are Javid and Gobustan. Define

A: Javid is ranked above Gobustan.

B: Javid is ranked first.

C: Javid is ranked 6th.

Give every probability to at least four significant figures.

(a) $P(A) =$ ____________

(b) $P(B) =$ ____________

(c) $P(A \cap B) =$ ____________

(d) $P(A \mid C) =$ ____________

(e) Are A and C independent?

- ○ Independent
- ○ Dependent

Solution

Nothing here needs the $9!$ orderings to be listed. Each part is settled by a symmetry argument on a small number of relevant positions.

(a) Consider only the two places occupied by Javid and Gobustan. By symmetry Javid is as likely to be the higher of the two as the lower, so $P(A) = 0.5$, whatever 9 is.

(b) Javid is equally likely to occupy each of the 9 places, so $P(B) = 1/9 = 0.1111$.

(c) If Javid is first it is above Gobustan automatically, so $B \subset A$ and $A \cap B = B$: $P(A \cap B) = 0.1111$.

(d) Suppose Javid is 6th. The other 8 funds fill the remaining places, and Gobustan is equally likely to be in any one of them. Of those places, $6 - 1$ are above Javid and $9 - 6 = 3$ are below, so

$$P(A \mid C) = \frac{3}{8} = 0.3750.$$

(e) That is different from [`P(A) = 0.5`], so [`A`] and [`C`] are dependent.

The general statement is worth extracting. With k funds and X in place j, $P(A \mid C) = (k - j)/(k - 1)$, and this equals $1/2$ exactly when $j = (k + 1)/2$ — the middle place. So the same event "X beats Y" is independent of "X finished in the middle" and dependent on every other placing. Independence is not a property you can see in the wording of two events; it is an arithmetic coincidence between their probabilities, and Definition 2.10 is the only way to find out whether it holds.

Problem 13

seed 20260923

A utility records whether each of its 462 business customers paid its last invoice late, by the region the customer is in.

	Late	On time	Total
Baku:	69	134	203
Ganja:	36	87	123
Sumqayit:	27	109	136

One customer is drawn at random. Let L be the event that it paid late and R_1, R_2, R_3 the events that it is in Baku, Ganja and Sumqayit.

Give every probability to at least four significant figures.

(a) $P(L) =$ ______________

(b) $P(L \mid R_2) =$ ______________

(c) $P(R_1 \mid L) =$ ______________

(d) $P(R_1 \cup R_2 \mid L) =$ ______________

Solution

(a) $69 + 36 + 27 = 132$ customers paid late, so $P(L) = 132/462 = 0.2857$.

(b) Among Ganja's 123 customers, 36 paid late: $P(L \mid R_2) = 36/123 = 0.2927$.

(c) Among the 132 late payers, 69 are in Baku: $P(R_1 \mid L) = 69/132 = 0.5227$.

(d) R_1 and R_2 are mutually exclusive, and conditioning does not change that, so the conditional probabilities add:

$$P(R_1 \cup R_2 \mid L) = P(R_1 \mid L) + P(R_2 \mid L) = \frac{69+36}{132} = \frac{105}{132} = 0.7955.$$

Worth noticing in general: for a fixed conditioning event L, the assignment $A \mapsto P(A \mid L)$ is itself a probability — it is non-negative, it gives the whole sample space probability one, and it adds over mutually exclusive events. Every rule proved for P therefore holds for $P(\cdot \mid L)$ as well, which is why part (d) needed no new argument.

Problem 14

seed 20260923

Two unrelated pieces of machinery in a plant break down independently in any given month. Let A be the event that the press breaks down and B that the furnace does, with $P(A) = 0.45$ and $P(B) = 0.29$.

Give every probability to at least four significant figures.

(a) $P(A \cap B) =$ ______________

(b) $P(A \cup B) =$ ______________

(c) A maintenance call-out is logged in any month in which at least one of the two breaks down. Given that a call-out was logged, what is the probability that the press broke down? ______________

(d) Given that a call-out was logged, what is the probability that the press broke down and the furnace did not?

____

Solution

(a) $P(A \cap B) = 0.45 \times 0.29 = 0.1305$ by independence.

(b) $P(A \cup B) = 0.45 + 0.29 - 0.1305 = 0.6095$.

(c) Write $U = A \cup B$. Since $A \subset U$, $A \cap U = A$ and

$$P(A \mid U) = \frac{P(A)}{P(U)} = \frac{0.45}{0.6095} = 0.7383.$$

(d) The event is $A \cap \bar{B}$, which is contained in U, and independence passes to complements:
$P(A \cap \bar{B}) = 0.45 \times (1 - 0.29)$. Dividing by $P(U)$,

$$P(A \cap \bar{B} \mid U) = 0.5242.$$

The moral of (c) and (d) is the one every maintenance log illustrates: the months on record are not a random sample of months. Conditioning on “something was logged” throws away the quiet months and inflates every breakdown rate computed from the log. Here the press’s unconditional rate is 0.45 and its rate among logged months is 0.7383, and neither figure is wrong — they answer different questions, and only the first is a property of the press.

Problem 15

seed 20260923

An auditor pulls three invoices at random, one after another and without replacement, from a batch of 61 invoices of which 46 are sound and 15 contain an error. Let S_1, S_2, S_3 be the events that the first, second and third invoice drawn is sound.

Give every probability to at least four significant figures.

(a) $P(S_2 \mid S_1) =$ ______________

(b) $P(S_3 \mid S_1 \cap S_2) =$ ______________

(c) $P(S_1 \cap S_2 \cap S_3) =$ ______________

(d) $P($at least one of the three contains an error$) =$ ______________

Solution

(a) After one sound invoice has been removed, 45 sound invoices remain among 60:
$P(S_2 \mid S_1) = 45/60 = 0.7500$.

(b) After two, $P(S_3 \mid S_1 \cap S_2) = 44/59 = 0.7458$.

(c) The multiplicative law applied twice:

$$P(S_1 \cap S_2 \cap S_3) = P(S_1)\, P(S_2 \mid S_1)\, P(S_3 \mid S_1 \cap S_2) = \frac{46}{61} \cdot \frac{45}{60} \cdot \frac{44}{59} = 0.4218.$$

(d) "At least one error" is the complement of "all three sound", so by Theorem 2.7 it is $1 - 0.4218 = 0.5782$.

Two remarks. First, each conditioning event in the chain is the **intersection** of everything drawn so far, not just the previous draw; writing $P(S_3 \mid S_2)$ would be a different and wrong quantity. Second, the chain could be written with any of the counting formulas of Section 2.6 instead — the number of ways to choose three sound invoices over the number of ways to choose any three — and the two expressions are equal. The sequential form is usually easier to write down and much easier to extend to four draws.

Problem 16

seed 20260923

Let B be the event that a registered firm is family owned and A the event that it has adopted electronic invoicing. A survey finds

$$P(B) = 0.69, \qquad P(A \mid B) = 0.45, \qquad P(A \mid \bar{B}) = 0.45.$$

Give every probability to at least four significant figures.

(a) $P(A \cap B) =$ ______________

(b) $P(A) =$ ______________

(c) $P(B \mid A) =$ ______________

(d) $P(A \cup B) =$ ______________

(e) Are A and B independent?

- ◯ Independent
- ◯ Dependent

Solution

(a) $P(A \cap B) = P(B)P(A \mid B) = 0.69 \times 0.45 = 0.3105$.

(b) Split A along B. The events $A \cap B$ and $A \cap \bar{B}$ are mutually exclusive and their union is A, so

$$P(A) = P(B)P(A \mid B) + P(\bar{B})P(A \mid \bar{B}) = 0.69 \times 0.45 + 0.3100 \times 0.45 = 0.45,$$

because the common factor 0.45 comes out and $P(B) + P(\bar{B}) = 1$.

(c) $P(B \mid A) = P(A \cap B)/P(A) = 0.3105/0.45 = 0.69$, which is $P(B)$ again.

(d) $P(A \cup B) = 0.45 + 0.69 - 0.3105 = 0.8295$.

(e) Independent. Part (b) is the general fact at work: if $P(A \mid B) = P(A \mid \bar{B})$, then both equal their weighted average, which is $P(A)$, and Definition 2.10 is satisfied.

This is the cleanest way to read what independence claims. Family firms adopt electronic invoicing at rate 0.45 and non-family firms at the same rate 0.45; ownership therefore carries no information about adoption, and the calculation in (c) shows the ignorance runs both ways. Independence is symmetric even though the phrase “A is independent of B” sounds as though it were not.

Problem 17

seed 20260923

A bank holds loans to three unrelated borrowers whose defaults over the coming year are independent events D_1, D_2, D_3 with

$$P(D_1) = 0.40, \qquad P(D_2) = 0.24, \qquad P(D_3) = 0.19.$$

Give every probability to at least four significant figures.

(a) $P(\text{no borrower defaults}) =$ ______________

(b) $P(\text{all three default}) =$ ______________

(c) $P(\text{exactly one defaults}) =$ ______________

(d) Given that exactly one borrower defaulted, what is the probability that it was borrower 1? ______________

Solution

Independence of D_1, D_2, D_3 lets every intersection be written as a product, and it carries over to the complements, so an event that specifies the fate of all three borrowers has probability equal to the product of three factors.

(a) $P(\bar{D}_1 \cap \bar{D}_2 \cap \bar{D}_3) = 0.3694$.

(b) $P(D_1 \cap D_2 \cap D_3) = 0.01824$.

(c) "Exactly one" is the union of three mutually exclusive outcomes — it must be borrower 1, or 2, or 3, and no two of these can happen at once — so the probabilities add:

$$0.24624 + 0.11664 + 0.08664 = 0.4495.$$

(d) Writing E for "exactly one defaulted", the event "borrower 1 defaults and the others do not" is contained in E, so

$$P(D_1 \mid E) = \frac{0.24624}{0.4495} = 0.5478.$$

The three answers in (d) — for borrowers 1, 2 and 3 — must add to one, which is a cheap and worthwhile check on the arithmetic.

Note how much the independence assumption is doing. It is what turns three marginal default probabilities into a full description of all eight possible outcomes. Borrowers in the same sector, or lending to each other, are not independent, and for them these three numbers do not determine the portfolio's risk at all — the probability that all three default can then be many times 0.01824. That gap is where a good deal of the 2008 losses lived.

Problem 18

seed 20260923

On a register of firms, A is the event that a firm exports and B the event that it holds a bank loan, with

$$P(A) = 0.65, \qquad P(B) = 0.47, \qquad P(A \cap B) = 0.2255.$$

Give every answer to at least four significant figures.

(a) What would $P(A \cap B)$ have to be for A and B to be independent? ______________

(b) $P(A \mid B) =$ ______________

(c) What would $P(A \mid B)$ have to be for A and B to be independent? ______________

Solution

(a) The third form of Definition 2.10 demands $P(A \cap B) = P(A)P(B) = 0.65 \times 0.47 = 0.3055$.

(b) From the observed intersection, $P(A \mid B) = 0.2255/0.47 = 0.4798$.

(c) The first form of Definition 2.10 demands $P(A \mid B) = P(A) = 0.65$.

The observed intersection 0.2255 is smaller than the 0.3055 that independence would require, and correspondingly $P(A \mid B) = 0.4798$ sits on the same side of $P(A) = 0.65$. The two comparisons always agree, because dividing both sides of $P(A \cap B)$ versus $P(A)P(B)$ by the positive number $P(B)$ preserves the inequality. Exporting and borrowing go together less often than independence would produce.

It is worth being careful about what this does and does not say. The calculation establishes association, not cause. Exporters may borrow to finance shipments, or banks may lend more readily to firms with foreign receivables, or a third characteristic — firm size, most obviously — may drive both. Nothing in Definition 2.10 distinguishes these stories, and a regulator who reads a departure from independence as a mechanism has gone beyond the arithmetic.

Problem 19

seed 20260923

A plant draws power through two generators feeding a single switchgear panel. The plant runs if **at least one** generator is working **and** the switchgear is working. The three components work independently, with probabilities 0.93, 0.80 and 0.94 respectively. Write G_1, G_2, W for the events that each works, and R for the event that the plant runs.

Give every probability to at least four significant figures.

(a) $P(G_1 \cup G_2) =$ ____________

(b) $P(R) =$ ____________

(c) $P(G_1 \mid R) =$ ____________

(d) $P(G_1 \cap G_2 \mid R) =$ ____________

Solution

(a) At least one generator is the complement of neither, and the two failures are independent:
$P(G_1 \cup G_2) = 1 - 0.0700 \times 0.2000 = 0.9860$.

(b) $R = (G_1 \cup G_2) \cap W$, and W is independent of the pair of generators, so $P(R) = 0.9860 \times 0.94 = 0.9268$.

(c) If generator 1 works, the parallel block works whatever generator 2 does, so $G_1 \cap R = G_1 \cap W$ and $P(G_1 \cap R) = 0.93 \times 0.94$. Dividing,

$$P(G_1 \mid R) = \frac{0.93 \times 0.94}{0.9268} = 0.9432.$$

(d) $(G_1 \cap G_2) \subset R$ once W is added, so the numerator is the triple product and

$$P(G_1 \cap G_2 \mid R) = \frac{0.93 \times 0.80 \times 0.94}{0.9268} = 0.7546.$$

Two things are worth taking away. First, the architecture matters more than the individual reliabilities: putting the generators in parallel lifts that block to 0.9860, above either generator on its own, while the switchgear sits in series and caps the whole plant at 0.94. A single series component is a ceiling no amount of redundancy elsewhere can raise.

Second, G_1 and G_2 are independent but **not** independent given R: part (d) is not the product of $P(G_1 \mid R)$ and $P(G_2 \mid R)$. Knowing the plant ran, a failure of one generator implies the other worked. Conditioning on a consequence of two independent causes makes them dependent.

Problem 20

seed 20260923

A fund has a track record of 10 years. In each year, independently of the others, it beats its benchmark with probability 0.345.

Give every probability to at least four significant figures.

(a) $P(\text{it never beat the benchmark}) =$ ______________

(b) $P(\text{it beat the benchmark in exactly one year}) =$ ______________

(c) Given that it beat the benchmark in exactly one year, what is the probability that the year in question was the first? ______________

(d) Given that it beat the benchmark in at least one year, what is the probability that it did so in the first year?

____

Solution

(a) Independence across years makes the probability a product of 10 identical factors: $0.6550^{10} = 0.01453$.

(b) Fix a year. The event "it beat the benchmark in that year and in no other" has probability $0.345 \times 0.6550^9 = 0.00766$, and this is the same for every year. The 10 such events are mutually exclusive, so

$$P(\text{exactly one}) = 10 \times 0.00766 = 0.0766.$$

(c) Each of the 10 one-success outcomes has the same probability, so conditioning on "exactly one" leaves them equally likely:

$$P(\text{the year was the first} \mid \text{exactly one}) = \frac{0.00766}{0.0766} = \frac{1}{10} = 0.1000.$$

Note what has dropped out: the answer does not depend on 0.345 at all. Once you know there was exactly one good year, the identity of that year is uniform over the 10 years, because nothing in the model distinguishes one year from another. Conditioning has removed the only quantity that carried the fund's skill.

(d) Now the conditioning event is L, "at least one", with $P(L) = 1 - 0.01453 = 0.9855$. The event "beat it in the first year" is contained in L, so

$$P(\text{first year} \mid L) = \frac{0.345}{0.9855} = 0.3501.$$

Parts (c) and (d) are the reason performance records are read badly. A fund that advertises "we beat the benchmark in year 3" is reporting the outcome of a search over 10 years, and under this model the **best** year of a fund with no skill at all is guaranteed to look like evidence of something. The probability of at least one good year, 0.9855, is the number to compare against, not the single-year probability 0.345.

Answer key

Problem	Answers
1	(a) 0.489362, (b) 0.784091, (c) 0.686567
2	(a) 0.89, (b) 0.644928, (c) 0.745, (d) Dependent
3	(a) 0.47037, (b) 0.556962, (c) 74.3185
4	(a) 0.617647, (b) 0.61194, (c) 0.377963, (d) 0.626866
5	(a) 0.231, (b) 0.679, (c) 0.641667
6	(a) 0.37, (b) 0.13, (c) 0.13, (d) 0.74
7	(a) 0.26, (b) 0.74, (c) 0, (d) Dependent, (e) 0.594595
8	(a) 0.825075, (b) 0.005075, (c) 0.994925, (d) 0.16985
9	(a) 0.2368, (b) 0.32, (c) 1, (d) 0.74, (e) 0.5032
10	(a) 0.5916, (b) 0.9584, (c) 0.907763, (d) 0.617279
11	(a) 0.0684932, (b) 0.0797753, (c) 0.71
12	(a) 0.5, (b) 0.111111, (c) 0.111111, (d) 0.375, (e) Dependent
13	(a) 0.285714, (b) 0.292683, (c) 0.522727, (d) 0.795455
14	(a) 0.1305, (b) 0.6095, (c) 0.73831, (d) 0.5242
15	(a) 0.75, (b) 0.745763, (c) 0.421784, (d) 0.578216
16	(a) 0.3105, (b) 0.45, (c) 0.69, (d) 0.8295, (e) Independent
17	(a) 0.36936, (b) 0.01824, (c) 0.44952, (d) 0.547784
18	(a) 0.3055, (b) 0.479787, (c) 0.65
19	(a) 0.986, (b) 0.92684, (c) 0.943205, (d) 0.754564
20	(a) 0.0145349, (b) 0.076558, (c) 0.1, (d) 0.350089