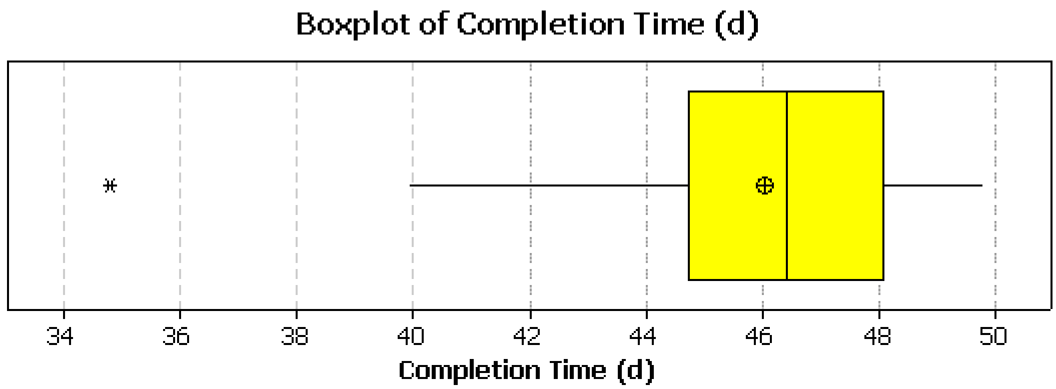
1) Observations of the times d (in days) for the completion of the same task by a sample of 54 contractors are summarized by the following Minitab® output:

Descriptive Statistics: Completion Time (d)

Variable	N	Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Completion Time (d)	54	46.049	2.711	34.789	44.752	46.415	48.057	49.791



(a) State any ***two*** reasons to conclude that this sample is negatively skewed. [2]
(b) Is the outlier mild or extreme? *Show your working*. [4]
(c) Estimate, to the nearest day, the second shortest completion time. [2]
(d) Comment briefly on the outlier - is it a plausibly genuine observation? Why or why not? [3]

(a) [Any two of]

- mean < median
- the left whisker is much longer than the right whisker
- the only outlier is far to the left [although see part (d) below]

(b) $IQR = Q3 - Q1 = 48.057 - 44.752 = 3.305$
Lower outer fence = $Q1 - 3IQR = 44.752 - 9.915 = 34.837$
The outlier is clearly the lowest value, at $x = 34.789$, just beyond the lower outer fence.
Therefore the outlier is **extreme**.

(c) x_2 is at the end of the left whisker. Therefore, to the nearest integer, $x_2 = \mathbf{40}$

(d) The outlier is extreme. For many distributions, extreme outliers occur by chance much less than 1% of the time. There are only 54 data here. The outlier is therefore unlikely to be a genuine value from the population from which the other 53 values were drawn. Measurement error (or fraud!) is a much more likely explanation for the extreme outlier. Therefore

NO, the outlier is suspect.

2) Events A, B, C form a partition. A bookmaker offers the following odds:
r_A = 5:3 on, r_B = 1:1 ("even odds") and r_C = 7:1 against

(a) Show that the corresponding probabilities are not coherent. [4]
(b) If ten deposits of $10 are placed with the quoted odds as follows: [5]
five deposits on event A, four on event B and one on event C;
then what is the bookmaker's profit (or loss) if event C occurs?
(c) Rescale the three probabilities so that they are coherent. [3]
(d) Convert the coherent probabilities back into odds. [2]

(a) $r_A = \frac{5}{3} \quad \Rightarrow \quad p_A = \frac{\frac{5}{3}}{\frac{5}{3}+1} = \frac{5}{5+3} = \frac{5}{8}$

$r_B = \frac{1}{1} \quad \Rightarrow \quad p_B = \frac{1}{1+1} = \frac{1}{2}$

$r_C = \frac{1}{7} \quad \Rightarrow \quad p_C = \frac{\frac{1}{7}}{\frac{1}{7}+1} = \frac{1}{1+7} = \frac{1}{8}$

$\Rightarrow \quad p_A + p_B + p_C = \frac{5}{8} + \frac{1}{2} + \frac{1}{8} = \frac{10}{8} = \frac{5}{4} > 1$

Therefore the "probabilities" are not coherent.

(b) bookmaker's profit = total revenue – total payout $= \sum k_i p_i s_i - k_C s_C$

$= \sum k_i b - k_C \frac{b}{p_C} = (5+4+1)\times 10 - 1\times \frac{10}{\frac{1}{8}} = 100 - 80 \quad \Rightarrow$

$\boxed{\text{profit} = \mathbf{\$20}}$

(c) $\sum_i p_i = \frac{5}{4}$ Therefore divide all three "probabilities" by $\frac{5}{4}$:

$p_A \to \frac{5}{8}\times\frac{4}{5} = \boxed{\frac{1}{2}}, \qquad p_B \to \frac{1}{2}\times\frac{4}{5} = \boxed{\frac{2}{5}}, \qquad p_C \to \frac{1}{8}\times\frac{4}{5} = \boxed{\frac{1}{10}}$

(d) $p_A = \frac{\frac{1}{2}}{1-\frac{1}{2}} = \frac{1}{2-1} = \boxed{1:1 \text{ (evens)}}, \qquad p_B = \frac{\frac{2}{5}}{1-\frac{2}{5}} = \frac{2}{5-2} = \boxed{3:2 \text{ against}},$

$p_C = \frac{\frac{1}{10}}{1-\frac{1}{10}} = \frac{1}{10-1} = \boxed{9:1 \text{ against}}$

[In part (b), note that the number of bids is in the same ratio as the probabilities, 5:4:1. In such situations, the bookmaker enjoys a profit of $\left(\sum \text{bids}\right)\times\left(1 - 1\Big/\sum_i p_i\right) = \$20$, no matter which event occurs!]

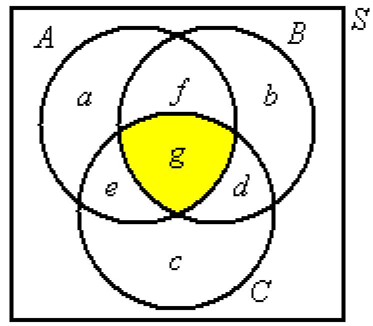
3) It is known that [8]

$P[A] = .40, \quad P[B] = .35, \quad P[C] = .50,$

$P[A \cup B] = .60, \quad P[B \cup C] = .75, \quad P[C \cup A] = .70 \quad \text{and} \quad P[A \cup B \cup C] = .85.$

Find the probability that ***all*** of events A, B, C occur.

$g = P[A \cap B \cap C]$ is required.



Start from the outside and work inwards.
Among the many valid methods is:

$a = P[A\tilde{B}\tilde{C}] = P[A \cup B \cup C] - P[B \cup C] = .85 - .75 = .10$

$b = P[\tilde{A}B\tilde{C}] = P[A \cup B \cup C] - P[A \cup C] = .85 - .70 = .15$

$c = P[\tilde{A}\tilde{B}C] = P[A \cup B \cup C] - P[A \cup B] = .85 - .60 = .25$

$b + d = P[\tilde{A}B] = P[A \cup B] - P[A] = .60 - .40 = .20$

$\Rightarrow \quad d = .20 - b = .20 - .15 = .05$

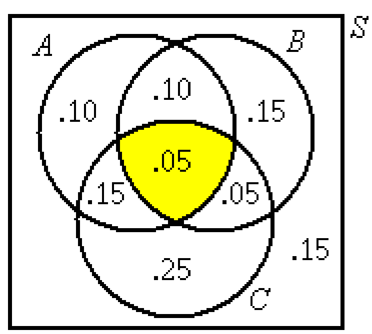
$a + e = P[A\tilde{B}] = P[A \cup B] - P[B] = .60 - .35 = .25$

$\Rightarrow \quad e = .25 - a = .25 - .10 = .15$

$P[C] = c + d + e + g \quad \Rightarrow \quad g = .50 - .25 - .05 - .15 = .05 \quad \Rightarrow$

$$P[A \cap B \cap C] = \mathbf{.05}$$

The complete Venn diagram [not essential] follows.



4) A truck is carrying fifteen coils of cables, three of which are defective.
As a random sample, four coils are removed from the truck.

(a) Find the probability that none of the four coils is defective. [4]
Express your answer as a fraction reduced to its lowest terms ***and*** as a decimal correct to two significant figures.

(b) Write down the probability mass function $p(x)$ for X, the number of defective coils in the random sample. [3]

BONUS QUESTION

(c) Using $\mathrm{E}[X] = \sum_{x=x_{\min}}^{x_{\max}} x \cdot p(x)$, find an exact expression for $\mathrm{E}[X]$. [+4]

(a) Let E represent the desired event, "none of the four coils is defective", then all four coils in the sample must be drawn from the (15 – 3 =) 12 non-defective coils.

$$\mathrm{P}[E] = \frac{n(E)}{n(S)} = \frac{{}^{12}C_4}{{}^{15}C_4} = \frac{12 \times 11 \times 10 \times 9}{4 \times 3 \times 2 \times 1} \times \frac{4 \times 3 \times 2 \times 1}{15 \times 14 \times 13 \times 12} = \frac{11 \times 3}{7 \times 13} \Rightarrow$$

$$\boxed{\mathrm{P}[E] = \frac{33}{91}}$$

(or $.\dot{3}6263\dot{7}$) Correct to two decimal places, P[E] = **<u>.36</u>** .

(b) In order for the random sample of size 4 to contain exactly x defective coils, x coils must be selected from the 3 defective coils **and** (n – x) coils must be selected from the 12 non-defective coils. Therefore

$$\boxed{p(x) = \mathrm{P}[X = x] = \frac{{}^{3}C_x \times {}^{12}C_{4-x}}{{}^{15}C_4}}$$

This expression is valid for all x, but produces non-zero values for x = 0, 1, 2, 3 only.
[Not required until part (c): The numerical values of this probability mass function are:
$p(0) = \frac{165}{455} = \frac{33}{91}$ (part (a)), $p(1) = \frac{220}{455} = \frac{44}{91}$, $p(2) = \frac{66}{455}$ and $p(3) = \frac{4}{455}$]

(c) $$\mathrm{E}[X] = \sum_{x=0}^{3} x \cdot p(x) = 0 + 1 \times \frac{220}{455} + 2 \times \frac{66}{455} + 3 \times \frac{4}{455} = \frac{364}{455} \Rightarrow$$

$$\boxed{\mu = \mathrm{E}[X] = \frac{4}{5}}$$

OR [next page]

To the continuation of this question ☞

4 (c) (continued)

The probability distribution here is hypergeometric, with population size $N = 15$, sample size $n = 4$, total number of successes (defective coils) in the population $R = 3$ and number of successes in the sample = x.
A more thorough and general derivation of $\mathrm{E}[X]$ is

$$\mathrm{E}[X] = \sum_x x \cdot p(x) = \sum_x \frac{x \cdot {}^R C_x \cdot {}^{N-R}C_{n-x}}{{}^N C_n}$$

But $x \cdot {}^R C_x = x \cdot \dfrac{R!}{x!(R-x)!} = x \cdot \dfrac{R(R-1)!}{x(x-1)!((R-1)-(x-1))!} = R \dfrac{S!}{y!(S-y)!} = R \cdot {}^S C_y$,

(where $S = R - 1$ and $y = x - 1$)

and ${}^N C_n = \dfrac{N!}{n!(N-n)!} = \dfrac{N(N-1)!}{n(n-1)!((N-1)-(n-1))!} = \dfrac{N}{n} \cdot \dfrac{M!}{m!(M-m)!} = \dfrac{N}{n} \cdot {}^M C_m$,

(where $M = N - 1$ and $m = n - 1$)

Also ${}^{N-R}C_{n-x} = {}^{(N-1)-(R-1)}C_{(n-1)-(x-1)} = {}^{M-S}C_{m-y}$

$$\Rightarrow \quad \mathrm{E}[X] = \frac{nR}{N} \sum_y \frac{{}^S C_y \cdot {}^{M-S}C_{m-y}}{{}^M C_m}$$

But the term inside this latter summation is the hypergeometric probability mass function for population size M, population number of successes S and sample size m. The sum is taken over all values of y. Because the p.m.f. is coherent, this sum must be $\sum_y p(y) = 1$. Therefore, for *any* hypergeometric distribution,

$$\boxed{\mu = \mathrm{E}[X] = \frac{nR}{N}}$$

With $N = 15$, $R = 3$ and $n = 4$, we obtain $\mathrm{E}[X] = \dfrac{4 \times 3}{15} = \dfrac{4}{5}$.

☜ Back to the index of solutions