Adding independent random variables

Starter

1. (Review of last lesson) A model is proposed for the weight distribution of trout in a fish farm. The density function of the model is $f(x) = \begin{cases} \frac{1}{36}w(6-w) & 0 \le x \le 6 \\ 0 & \text{otherwise} \end{cases}$.

A sample of 80 fish is obtained and their weights are recorded. The table contains a summary of the data.

Weight class	$0 \le w < 1$	$1 \le w < 2$	$2 \le w < 4$	$4 \le w < 6$
Number of fish	6	14	39	21

(a) Use the model to determine the expected number of fish in each weight class.
(b) Conduct a goodness-of-fit test at the $10\,\%$ level to determine if the model is appropriate for these data.

Working: (a) The expected number of fish in each weight class are:

$$0 \le w < 1: \quad 80 \times \int_0^1 \frac{1}{36}w(6-w)dw = \frac{160}{27} \approx 5.93$$

$$1 \le w < 2: \quad 80 \times \int_1^2 \frac{1}{36}w(6-w)dw = \frac{400}{27} \approx 14.8$$

$$2 \le w < 4: \quad 80 \times \int_2^4 \frac{1}{36}w(6-w)dw = \frac{1040}{27} \approx 38.5$$

$$4 \le w < 6: \quad 80 \times \int_4^6 \frac{1}{36}w(6-w)dw = \frac{560}{27} \approx 20.7$$

(b) $$\chi^2_{calc} = \sum \frac{(O_i - E_i)^2}{E_i}$$

$$= \frac{(6-5.93)^2}{5.93} + \frac{(14-14.8)^2}{14.8} + \frac{(39-38.5)^2}{38.5} + \frac{(21-20.8)^2}{20.8}$$

$\chi^2_{calc} \approx 0.0549$

Degrees of freedom, $\nu = 4 - 1 = 3$

The critical value at the $10\,\%$ level is $\chi^2_3(10\%) = 6.251$

H_0 : the data can be modelled by the proposed model

H_1 : the data cannot be modelled by the proposed model

Since $\chi^2_{calc} \approx 0.0549 < 6.251 = \chi^2_3(5\%)$, we do not reject H_0.

There is evidence to suggest that the proposed model is a good fit.

E.g. 1 From AS level we have $\mathrm{E}(aX+b) = a\mathrm{E}(X)+b$ and $\text{Var}(aX+b) = a^2\text{Var}(X)$. Given that if X and Y are independent random variables then $E(X+Y) = E(X)+E(Y)$ and $\text{Var}(X+Y) = \text{Var}(X)+\text{Var}(Y)$, conjecture expressions for the following:

(a) $E(X-Y)$
(b) $\text{Var}(X-Y)$
(c) $E(aX+bY+c) = aE(X)+bE(Y)+c$
(d) $\text{Var}(aX+bY+c) = a^2\text{Var}(X)+b^2\text{Var}(Y)$

Working: (a) $E(X-Y) = E(X+(-Y)) = E(X)-E(Y)$

(b) $$\begin{aligned}\text{Var}(X-Y) &= \text{Var}(X+(-Y))\\ &= \text{Var}(X)+(-1)^2\text{Var}(Y)\\ &= \text{Var}(X)+\text{Var}(Y)\end{aligned}$$

(c) $E(aX+bY+c) = aE(X)+bE(Y)+c$

(d) $\text{Var}(aX+bY+c) = a^2\text{Var}(X)+b^2\text{Var}(Y)$

E.g. 2 Independent random variables X and Y are such that $E(X)=3$, $E(Y)=5$, $\text{Var}(X)=4$ and $\text{Var}(Y)=2$. Find:

(a) $E(4X+2Y)$
(b) $E(6X-Y)$
(c) $\text{Var}(3X+6Y+11)$
(d) $\text{Var}(5Y-3X)$
(e) $\text{Var}(3X-5Y)$

Working: (a) $E(4X+2Y) = 4\times3+2\times5 = 22$

(b) $E(6X-Y) = 6\times3-5 = 13$

(c) $\text{Var}(3X+6Y+11) = 3^2\times4+6^2\times2 = 108$

(d) $\text{Var}(5Y-3X) = 5^2\times2+3^2\times4 = 86$

(e) $\text{Var}(3X-5Y) = 3^2\times4+5^2\times2 = 86$

E.g. 3 Let X_1 and X_2 be independent observations of the random variable X so that:

$$E(X_1) = E(X_2) = E(X) \qquad \text{and} \qquad \text{Var}(X_1) = \text{Var}(X_2) = \text{Var}(X)$$

Show that

(a) $E(X_1+X_2) = E(2X)$
(b) $\text{Var}(X_1+X_2) \neq \text{Var}(2X)$?

Working: (a) $$\begin{aligned}E(X_1+X_2) &= E(X_1)+E(X_2)\\ &= E(X)+E(X)\\ &= 2E(X)\\ &= E(2X)\end{aligned}$$

(b) $\text{Var}(X_1+X_2) = \text{Var}(X_1)+\text{Var}(X_2) = \text{Var}(X)+\text{Var}(X) = 2\text{Var}(X)$
But $\text{Var}(2X) = 2^2\text{Var}(X) = 4\text{Var}(X)$
In fact, $\text{Var}(2X) > \text{Var}(X_1)+\text{Var}(X_2)$

E.g. 4 A crane is lifting a crate with 4 large boxes and 5 small boxes. The large boxes have mean mass 18 kg and standard deviation 3 kg while the small boxes have mean mass 12 kg and standard deviation 1.5 kg. Given that the crate has mass 25 kg, calculate the expectation and standard deviation of the total mass of the crate with 4 large boxes and 5 small boxes loaded on it.

Working: $E(L) = 18$ and $\text{Var}(L) = 3^2$
$E(S) = 12$ and $\text{Var}(S) = 1.5^2$

$$\begin{aligned} E(4 \text{ large \& } 5 \text{ small boxes plus crate}) &= E(L_1) + \ldots + E(L_4) + E(S_1) + \ldots + E(S_5) + 25 \\ &= 4E(L) + 5E(S) \\ &= 4 \times 18 + 5 \times 12 + 25 \\ &= 157 \end{aligned}$$

$$\begin{aligned} \text{Var}(4 \text{ large \& } 5 \text{ small boxes plus crate}) &= \text{Var}(L_1) + \ldots + \text{Var}(L_4) + \text{Var}(S_1) + \ldots + \text{Var}(S_5) \\ &= 4\text{Var}(L) + 5\text{Var}(S) \\ &= 4 \times 3^2 + 5 \times 1.5^2 \\ &= 47.25 \end{aligned}$$

The expectation and standard deviation are 157 and $\dfrac{3\sqrt{21}}{2} = 6.87$.

E.g. 5 Two Year 7 tutor groups, A and B, take a maths test and the details of their results are:

	Mean	**Standard deviation**
$7A$	69	8.1
$7B$	75	5.2

One student from $7A$ and one student from $7B$ are selected at random. Calculate the expected mean and standard deviation of the difference between their scores.

Working: $E(\text{difference between scores}) = E(7B) - E(7A) = 6$

$$\begin{aligned} \text{Var}(\text{difference between scores}) &= \text{Var}(7A) + \text{Var}(7B) \\ &= 8.1^2 + 5.2^2 \\ &= 92.65 \end{aligned}$$

Video: **Combining independent random variables**

Solutions to Starter and E.g.s

Exercise
p153 8A Qu 1i, 3-6, (7 red)