Nice properties of $E[g(\boldsymbol{X})]$, $E[\alpha\boldsymbol{X} + \beta]$, $var(\boldsymbol{X})$ and $var(\alpha\boldsymbol{X} + \beta)$

BIO210 Biostatistics

Extra reading material for Lecture 10

Xi Chen
School of Life Sciences
Southern University of Science and Technology

March 2022

1 About $E[g(\boldsymbol{X})]$

A ***random variable*** is a ***function*** that maps an outcome from the ***sample space*** to a real-valued number. Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be discrete random variables (*r.v.*), and

$$\boldsymbol{Y} = g(\boldsymbol{X})$$

Then

$$E[\boldsymbol{Y}] = \sum_x g(x) p_{\boldsymbol{X}}(x)$$

Where $p_{\boldsymbol{X}}(x)$ is the ***probability mass function*** (PMF) of the *r.v.* $\boldsymbol{X}$.

Proof

Proof. According to the definition of *expectation*, we have

$$E[\boldsymbol{Y}] = \sum_y y p_{\boldsymbol{Y}}(y) \tag{1}$$

Note the relationship between *r.v.* $\boldsymbol{X}$ and $\boldsymbol{Y}$, we have

$$p_{\boldsymbol{Y}}(y) = \sum_{x \mid g(x)=y} p_{\boldsymbol{X}}(x) \tag{2}$$

Put the equation (2) into (1), we have

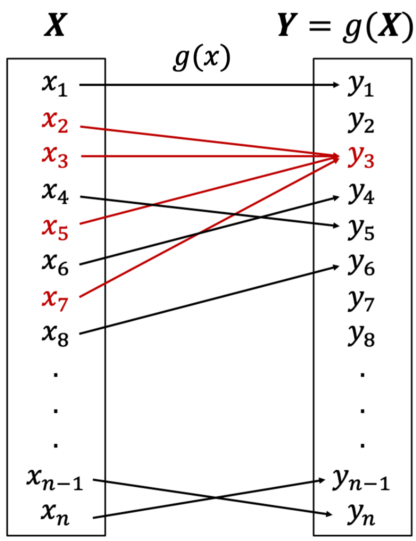
$$\begin{aligned} p_{\boldsymbol{Y}}(y) &= \sum_y y \sum_{x \mid g(x)=y} p_{\boldsymbol{X}}(x) \\ &= \sum_y \sum_{x \mid g(x)=y} y p_{\boldsymbol{X}}(x) \\ &= \sum_y \sum_{x \mid g(x)=y} g(x) p_{\boldsymbol{X}}(x) \\ &= \sum_x g(x) p_{\boldsymbol{X}}(x) \end{aligned}$$

□

If you are not comfortable with the $\sum$ notation, you can look at the diagram at the right. Since

$$E[\boldsymbol{Y}] = \sum_y y p_{\boldsymbol{Y}}(y)$$

We need to calculate every single $y_i p_{\boldsymbol{Y}}(y_i)$ in order to get $E[\boldsymbol{Y}]$. Here, we use the coloured (red) example as a demonstration.



In this case, we are calculating $y_3 \cdot p_{\boldsymbol{Y}}(\boldsymbol{Y} = y_3)$. Note $p_{\boldsymbol{Y}}(\boldsymbol{Y} = y_3)$ is the sum of all x_i such that $g(x_i) = y_3$. In this specific example, they are $\{x_2, x_3, x_5, x_7\}$. Therefore, we have:

$$\begin{aligned} y_3.p_{\boldsymbol{Y}}(\boldsymbol{Y} = y_3) &= y_3.[p_{\boldsymbol{X}}(x_2) + p_{\boldsymbol{X}}(x_3) + p_{\boldsymbol{X}}(x_5) + p_{\boldsymbol{X}}(x_7)] \\ &= y_3.p_{\boldsymbol{X}}(x_2) + y_3.p_{\boldsymbol{X}}(x_3) + y_3.p_{\boldsymbol{X}}(x_5) + y_3.p_{\boldsymbol{X}}(x_7) \end{aligned}$$

Note $y_3 = g(x_2) = g(x_3) = g(x_5) = g(x_7)$, then the above equation becomes:

$$\begin{aligned} y_3.p_{\boldsymbol{Y}}(\boldsymbol{Y} = y_3) &= g(x_2)p_{\boldsymbol{X}}(x_2) + g(x_3)p_{\boldsymbol{X}}(x_3) + g(x_5)p_{\boldsymbol{X}}(x_5) + g(x_7)p_{\boldsymbol{X}}(x_7) \\ &= \sum_{i \in \{2,3,5,7\}} g(x_i)p_{\boldsymbol{X}}(x_i) \end{aligned}$$

2 About $E[\alpha\boldsymbol{X} + \beta]$

The linear function $\alpha\boldsymbol{X} + \beta$ scale the *r.v.* $\boldsymbol{X}$ by a constant factor α and shift everything by a constant factor β. Therefore, we intuitively should expect that:

$$E[\alpha\boldsymbol{X} + \beta] = \alpha E[\boldsymbol{X}] + \beta$$

Proof

Proof. Using the property that $E[g(\boldsymbol{X})] = \sum_x g(x)p_{\boldsymbol{X}}(x)$, we have:

$$\begin{aligned} E[\alpha\boldsymbol{X} + \beta] &= \sum_x (\alpha x + \beta)p_{\boldsymbol{X}}(x) = \sum_x [\alpha x p_{\boldsymbol{X}}(x) + \beta p_{\boldsymbol{X}}(x)] \\ &= \sum_x \alpha x p_{\boldsymbol{X}}(x) + \sum_x \beta p_{\boldsymbol{X}}(x) \end{aligned}$$

Since α and β are constants, we could take them out from the summation:

$$E[\alpha\boldsymbol{X} + \beta] = \alpha \sum_x x p_{\boldsymbol{X}}(x) + \beta \sum_x p_{\boldsymbol{X}}(x)$$

Note that by definition, $\sum_x x p_{\boldsymbol{X}}(x) = E[\boldsymbol{X}]$ and $\sum_x p_{\boldsymbol{X}}(x) = 1$, we have:

$$E[\alpha\boldsymbol{X} + \beta] = \alpha E[\boldsymbol{X}] + \beta$$

□

3 About $var(\boldsymbol{X})$

The definition is:

$$var(\boldsymbol{X}) = E[(\boldsymbol{X} - E[\boldsymbol{X}])^2]$$

Let's look at this bit by bit. First, for any given *r.v.* $\boldsymbol{X}$, $E[\boldsymbol{X}]$ is a constant value. Then $(\boldsymbol{X} - E[\boldsymbol{X}])^2$ is a function of the *r.v.* $\boldsymbol{X}$, so $(\boldsymbol{X} - E[\boldsymbol{X}])^2$ is also an *r.v.*. Therefore, it is reasonable to ask: what is the expectation of $(\boldsymbol{X} - E[\boldsymbol{X}])^2$? This is basically $E[(\boldsymbol{X} - E[\boldsymbol{X}])^2]$, and it represents how much $\boldsymbol{X}$ deviates from its mean. Very often, it is actually easier to calculate the variance using the following formula:

$$var(\boldsymbol{X}) = E[\boldsymbol{X}^2] - (E[\boldsymbol{X}])^2$$

Proof

Proof. We start from the definition, use the property of $E[g(\boldsymbol{X})]$, and expand the thing inside the parentheses:

$$\begin{aligned} var(\boldsymbol{X}) &= E[(\boldsymbol{X} - E[\boldsymbol{X}])^2] = \sum_x (x - E[\boldsymbol{X}])^2 p_{\boldsymbol{X}}(x) \\ &= \sum_x \left[x^2 p_{\boldsymbol{X}}(x) - 2xE[\boldsymbol{X}]p_{\boldsymbol{X}}(x) + (E[\boldsymbol{X}])^2 p_{\boldsymbol{X}}(x)\right] \\ &= \sum_x x^2 p_{\boldsymbol{X}}(x) - \sum_x 2xE[\boldsymbol{X}]p_{\boldsymbol{X}}(x) + \sum_x (E[\boldsymbol{X}])^2 p_{\boldsymbol{X}}(x) \end{aligned}$$

Since $E[\boldsymbol{X}]$ is a constant value, we can take it out of the summation:

$$var(\boldsymbol{X}) = \sum_x x^2 p_{\boldsymbol{X}}(x) - 2E[\boldsymbol{X}] \sum_x x p_{\boldsymbol{X}}(x) + (E[\boldsymbol{X}])^2 \sum_x p_{\boldsymbol{X}}(x)$$

Now, note that $\sum_x x^2 p_{\boldsymbol{X}}(x) = E[\boldsymbol{X}^2]$ and $\sum_x x p_{\boldsymbol{X}}(x) = E[\boldsymbol{X}]$ by definition, and $\sum_x p_{\boldsymbol{X}}(x) = 1$. Therefore, we have:

$$\begin{aligned} var(\boldsymbol{X}) &= E[\boldsymbol{X}^2] - 2E[\boldsymbol{X}] \cdot E[\boldsymbol{X}] + (E[\boldsymbol{X}])^2 \\ &= E[\boldsymbol{X}^2] - 2(E[\boldsymbol{X}])^2 + (E[\boldsymbol{X}])^2 \\ &= E[\boldsymbol{X}^2] - (E[\boldsymbol{X}])^2 \end{aligned}$$

□

4 About $var(\alpha\boldsymbol{X} + \beta)$

Again, $\alpha\boldsymbol{X} + \beta$ means scaling the *r.v.* $\boldsymbol{X}$ by a constant factor α and shifting everything by a constant factor β. When scaling the values, the scaling factor will be exaggerated by the square operation in the variance formula. When shifting the values, the shape of the distribution does not change, so the variance will not be affected by the shifting factor. We should not be surprised that:

$$var(\alpha \boldsymbol{X} + \beta) = \alpha^2 var(\boldsymbol{X})$$

Proof

Proof. We can start with the definition:

$$var(\alpha \boldsymbol{X} + \beta) = E[(\ (\alpha \boldsymbol{X} + \beta) - E(\alpha \boldsymbol{X} + \beta)\ )^2]$$

Note that $E(\alpha \boldsymbol{X} + \beta) = \alpha \boldsymbol{X} + \beta$, so we have:

$$\begin{aligned}
var(\alpha \boldsymbol{X} + \beta) &= E[(\ (\alpha \boldsymbol{X} + \beta) - (\alpha E[\boldsymbol{X}] + \beta)\ )^2] \\
&= E[(\alpha \boldsymbol{X} + \beta - \alpha E[\boldsymbol{X}] - \beta)^2] \\
&= E[(\alpha \boldsymbol{X} - \alpha E[\boldsymbol{X}])^2] \\
&= E[\alpha^2 \boldsymbol{X}^2 - 2\alpha^2 \boldsymbol{X} E[\boldsymbol{X}] + \alpha^2 (E[\boldsymbol{X}])^2] \\
&= \sum_x [\alpha^2 x^2 - 2\alpha^2 x E[\boldsymbol{X}] + \alpha^2 (E[\boldsymbol{X}])^2] p_{\boldsymbol{X}}(x) \\
&= \sum_x \alpha^2 x^2 p_{\boldsymbol{X}}(x) - \sum_x 2\alpha^2 x E[\boldsymbol{X}] p_{\boldsymbol{X}}(x) + \sum_x \alpha^2 (E[\boldsymbol{X}])^2 p_{\boldsymbol{X}}(x) \\
&= \alpha^2 \sum_x x^2 p_{\boldsymbol{X}}(x) - 2\alpha^2 E[\boldsymbol{X}] \sum_x x p_{\boldsymbol{X}}(x) + \alpha^2 (E[\boldsymbol{X}])^2 \sum_x p_{\boldsymbol{X}}(x)
\end{aligned}$$

Again, we already know that $\sum_x x^2 p_{\boldsymbol{X}}(x) = E[\boldsymbol{X}^2]$, $\sum_x x p_{\boldsymbol{X}}(x) = E[\boldsymbol{X}]$ and $\sum_x p_{\boldsymbol{X}}(x) = 1$. Therefore:

$$\begin{aligned}
var(\alpha \boldsymbol{X} + \beta) &= \alpha^2 E[\boldsymbol{X}^2] - 2\alpha^2 E[\boldsymbol{X}] \cdot E[\boldsymbol{X}] + \alpha^2 (E[\boldsymbol{X}])^2 \\
&= \alpha^2 E[\boldsymbol{X}^2] - 2\alpha^2 (E[\boldsymbol{X}])^2 + \alpha^2 (E[\boldsymbol{X}])^2 \\
&= \alpha^2 E[\boldsymbol{X}^2] - \alpha^2 (E[\boldsymbol{X}])^2 \\
&= \alpha^2 (E[\boldsymbol{X}^2] - (E[\boldsymbol{X}])^2)
\end{aligned}$$

Remember we just proved that $var(\boldsymbol{X}) = E[\boldsymbol{X}^2] - (E[\boldsymbol{X}])^2$, so we have:

$$var(\alpha\boldsymbol{X} + \beta) = \alpha^2 var(\boldsymbol{X})$$

□

To be honest, all those proofs will be much easier if we know ***Linearity of Expectation***, which will be covered in the next lecture.