Linear Functions of Normal Random Variables

BIO210 Biostatistics

Extra reading material for Lecture 13

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

March 2022

1 A Linear Function of A Continuous Random Variable In General

Let's first have a look what happens to a ***random variable*** in general when we apply a linear function to it. Let $\boldsymbol{X}$ be a continuous random variable with a PDF $f_X(x)$. Let the random variable $\boldsymbol{Y}$ be:

$$\boldsymbol{Y} = a\boldsymbol{X} + b$$

where $a \neq 0$. What is the PDF of $\boldsymbol{Y}$?

Again, we should start with something simple. Consider this: if $\boldsymbol{X}$ and $\boldsymbol{Y}$ were discrete random variables, the situation becomes straightforward. We would have:

$$p_{\boldsymbol{Y}}(y) = P(\boldsymbol{Y} = y) = P(a\boldsymbol{X} + b = y) = P\left(\boldsymbol{X} = \frac{y-b}{a}\right)$$

However, for continuous random variables, the probability of getting a specific value is 0. Therefore, it is not very helpful to use the strategy above. We need to work on intervals for continuous random variables. The trick here is to use the CDF to solve the problem

1.1 When $a > 0$

Consider the case where $a > 0$, we have:

$$\begin{aligned} F_{\boldsymbol{Y}}(y) &= P(\boldsymbol{Y} \leqslant y) = P(a\boldsymbol{X} + b \leqslant y) \\ &= P\left(\boldsymbol{X} \leqslant \frac{y-b}{a}\right) = F_{\boldsymbol{X}}(x)\left(\frac{y-b}{a}\right) \end{aligned}$$

Now, that tells us the CDF of $\boldsymbol{Y}$ in terms of the CDF of $\boldsymbol{X}$:

$$F_{\boldsymbol{Y}}(y) = F_{\boldsymbol{X}}(x)\left(\frac{y-b}{a}\right)$$

Since the derivate of the CDF is the PDF, now we can simply find out the PDF by differentiating both sides of the above equation like this:

$$f_{\boldsymbol{Y}}(y) = F'_{\boldsymbol{X}}\left(\frac{y-b}{a}\right) = \frac{1}{a} \cdot f_{\boldsymbol{X}}\left(\frac{y-b}{a}\right) \tag{1}$$

1.2 When $a < 0$

Now consider the case where $a < 0$. Using the similar technique, we have:

$$\begin{aligned} F_{\boldsymbol{Y}}(y) &= P(\boldsymbol{Y} \leqslant y) = P(a\boldsymbol{X} + b \leqslant y) \\ &= P\left(\boldsymbol{X} \geqslant \frac{y-b}{a}\right) = 1 - P\left(\boldsymbol{X} \leqslant \frac{y-b}{a}\right) \\ &= 1 - F_{\boldsymbol{X}}(x)\left(\frac{y-b}{a}\right) \end{aligned}$$

Taking the derivate at the both sides of the above equation, we have:

$$f_{\boldsymbol{Y}}(y) = -F'_{\boldsymbol{X}}\left(\frac{y-b}{a}\right) = -\frac{1}{a} \cdot f_{\boldsymbol{X}}\left(\frac{y-b}{a}\right) \tag{2}$$

Combine the cases where $a > 0$ (1) and $a < 0$ (2), we have:

$$f_{\boldsymbol{Y}}(y) = \frac{1}{|a|} \cdot f_{\boldsymbol{X}}\left(\frac{y-b}{a}\right) \tag{3}$$

2 A Linear Function of A Normal Random Variable

Now, consider the normal random variable $\boldsymbol{X} \sim \mathcal{N}(\mu, \sigma^2)$, what is the PDF of the random variable $\boldsymbol{Y} = a\boldsymbol{X} + b$? We are given that:

$$f_{\boldsymbol{X}}(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Using equation (3) from the previous section, we have:

$$\begin{aligned} f_{\boldsymbol{Y}}(y) &= \frac{1}{|a|} \cdot f_{\boldsymbol{X}}\left(\frac{y-b}{a}\right) \\ &= \frac{1}{|a|} \cdot \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\frac{y-b}{a}-\mu)^2}{2\sigma^2}} \\ &= \frac{1}{\sqrt{2\pi}\sigma|a|} e^{-\frac{(y-b-a\mu)^2}{2a^2\sigma^2}} \end{aligned} \tag{4}$$

Re-write equation (4) a bit, we have:

$$f_{\boldsymbol{Y}}(y) = \frac{1}{\sqrt{2\pi} \cdot |a|\sigma} e^{-\frac{[y-(a\mu+b)]^2}{2(|a|\sigma)^2}} \tag{5}$$

From equation (5), we can easily see that $\boldsymbol{Y} \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$.