STAB52, An Introduction to Probability - Tutorial 4

Chong Chen

University of Toronto, Scarborough

Definition: Given a random variable X, it's *cumulative distribution function* is the function $F_X : \mathbb{R} \to [0,1]$ defined by

$$F_X(x) = \mathbb{P}(X \leq x).$$

One particularly special example is the normal distribution...

Problem 1. (The Normal Distribution)
Suppose that the grades of students in a given class are approximately normally distributed with mean 75 and variance 100. Find the probability of students who scored between 71 and 83.

Solution: Let X be the grade of a student in the class (a random variable). Then

$$\begin{aligned}\mathbb{P}(71 \leq X \leq 83) &= \int_{71}^{83} \frac{1}{\sqrt{200\pi}} e^{-\frac{1}{2}\left(\frac{x-75}{10}\right)^2} dx \\ &= \int_{-0.4}^{0.8} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du \\ &= \int_{-\infty}^{0.8} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du - \int_{-\infty}^{-0.4} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du \\ &= \Phi(0.8) - \Phi(-0.4) \\ &= 0.7881 - 0.3446 = 0.4435.\end{aligned}$$

Here, we have made the substitution $u = (x-75)/10$. We remark that unlike most absolutely continuous random variables, we cannot actually integrate the density of the standard normal distribution since there is no antiderivative for $e^{u^2/2}$. Instead, these cumulative density values, $\Phi(x)$ (z-scores) have been approximated using computer software and included in the appendix of the textbook. When we discuss the central limit theorem, the significance of normal distributions will become clear.

$\square$

Problem 2. (Evans & Rosenthal, Exercise 2.5.16.)
Let F be a cumulative distribution function. Compute (with explanation) the value of $\lim_{n\to\infty}[F(2n)-F(n)]$.

Solution: We compute

$$\lim_{n\to\infty}[F(2n) - F(n)] = \lim_{n\to\infty} F(2n) - \lim_{n\to\infty} F(n) = 1 - 1 = 0.$$

One must note that we may distribute the limit only because the limit of each term exists (this is a calculus result). Alternatively, we may approach this problem using continuity of probability:

$$\lim_{n\to\infty}[F(2n) - F(n)] = \lim_{n\to\infty} \mathbb{P}(n \leq X \leq 2n) \leq \lim_{n\to\infty} \mathbb{P}(n \leq X < \infty) = \mathbb{P}(\emptyset)$$

since $A_n := \{X \in [n,\infty)\} \searrow \emptyset$.

$\square$

Problem 3. (Evans & Rosenthal, Exercise 2.5.17.)
Let F be a cumulative distribution function. For $x \in \mathbb{R}$, we could define $F(x^+)$ by $F(x^+) = \lim_{n\to\infty} F(x+\frac{1}{n})$. Prove that F is *right continuous*, meaning that for each $x \in \mathbb{R}$, we have $F(x^+) = F(x)$.

Solution: First fix an arbitrary $x \in \mathbb{R}$. Let X be a random variable whose cumulative distribution function is F. Let $A_n = \{X \le x + \frac{1}{n}\}$ so then $A_n \searrow \{X \le x\}$. We then have by continuity of measure and the definition of the cumulative distribution function that

$$F(x) = \mathbb{P}(X \le x) = \lim_{n\to\infty} \mathbb{P}\left(X \le x + \frac{1}{n}\right) = \lim_{n\to\infty} F\left(x + \frac{1}{n}\right) = F(x^+)$$

$\square$

Problem 4. (Evans & Rosenthal, Exercise 2.5.18.)
Let X be a random variable, with cumulative distribution function F_X. Prove that $\mathbb{P}(X = a) = 0$ if and only if the function F_X is continuous at a.

Solution: For the first direction of implication, we assume that F_X is continuous at a. In particular, this means that F_X is left continuous i.e. $\lim_{n\to\infty} \mathbb{P}(X \le a - \frac{1}{n}) = \mathbb{P}(X \le a)$. On the other hand, if we let $A_n = \{X \le a - \frac{1}{n})$ so that $A_n \nearrow \{X < a\}$, then we have $\lim_{n\to\infty} \mathbb{P}(X \le a - \frac{1}{n}) = \mathbb{P}(X < a)$ by continuity of probability. Combining these results, we conclude $\mathbb{P}(X \le a) = \mathbb{P}(X < a)$ and hence $\mathbb{P}(X = a) = 0$ by disjoint additivity.

For the reverse direction, suppose $\mathbb{P}(X = a) = 0$. Then by the same continuity argument used above, we compute

$$\lim_{n\to\infty} \mathbb{P}(X \le a - \frac{1}{n}) = \mathbb{P}(X < a) = \mathbb{P}(X < a) + 0 = \mathbb{P}(X < a) + \mathbb{P}(X = a) = \mathbb{P}(X \le a).$$

This shows that F_X is left continuous at a. But the previous problem showed that F_X is right continuous and so F_X is continuous at a.

$\square$