1 First Exponential to Die

Let X and Y be Exponential(λ_1) and Exponential(λ_2) respectively, independent. What is

$$\mathbb{P}\big(\min(X,Y) = X\big),$$

the probability that the first of the two to die is X?

Solution:

Recall that the CDF of an exponential is $\mathbb{P}[X \leq x] = 1 - \exp(-\lambda x)$ for $x \geq 0$.

$$\mathbb{P}\big(\min(X,Y) = X\big) = \mathbb{P}(Y > X) = \int_0^\infty \mathbb{P}(Y > X \mid X = x) f_X(x)\,\mathrm{d}x = \int_0^\infty \mathrm{e}^{-\lambda_2 x} \cdot \lambda_1 \,\mathrm{e}^{-\lambda_1 x}\,\mathrm{d}x$$
$$= -\frac{\lambda_1}{\lambda_1 + \lambda_2}\mathrm{e}^{-(\lambda_1+\lambda_2)x}\Big|_{x=0}^\infty = \frac{\lambda_1}{\lambda_1+\lambda_2}.$$

2 Chebyshev's Inequality vs. Central Limit Theorem

Let n be a positive integer. Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with the following distribution:

$$\mathbb{P}[X_i = -1] = \frac{1}{12}; \qquad \mathbb{P}[X_i = 1] = \frac{9}{12}; \qquad \mathbb{P}[X_i = 2] = \frac{2}{12}.$$

(a) Calculate the expectations and variances of X_1, $\sum_{i=1}^n X_i$, $\sum_{i=1}^n (X_i - \mathbb{E}[X_i])$, and

$$Z_n = \frac{\sum_{i=1}^n (X_i - \mathbb{E}[X_i])}{\sqrt{n/2}}.$$

(b) Use Chebyshev's Inequality to find an upper bound b for $\mathbb{P}[|Z_n| \geq 2]$.

(c) Can you use b to bound $\mathbb{P}[Z_n \geq 2]$ and $\mathbb{P}[Z_n \leq -2]$?

(d) As $n \to \infty$, what is the distribution of Z_n?

(e) We know that if $Z \sim \mathcal{N}(0,1)$, then $\mathbb{P}[|Z| \leq 2] = \Phi(2) - \Phi(-2) \approx 0.9545$. As $n \to \infty$, can you provide approximations for $\mathbb{P}[Z_n \geq 2]$ and $\mathbb{P}[Z_n \leq -2]$?

Solution:

(a) $\mathbb{E}[X_1] = -1/12 + 9/12 + 4/12 = 1$, and

$$\operatorname{Var} X_1 = \frac{1}{12} \cdot 2^2 + \frac{9}{12} \cdot 0^2 + \frac{2}{12} \cdot 1^2 = \frac{1}{2}.$$

Using linearity of expectation and variance (since $X_1, \ldots, X_n$ are independent), we find that $\mathbb{E}[\sum_{i=1}^n X_i] = n$ and $\operatorname{var}(\sum_{i=1}^n X_i) = n/2$.

Again, by linearity of expectation, $\mathbb{E}[\sum_{i=1}^n X_i - n] = n - n = 0$. Subtracting a constant does not change the variance, so $\operatorname{var}(\sum_{i=1}^n X_i - n) = n/2$, as before.

Using the scaling properties of the expectation and variance, $\mathbb{E}[Z_n] = 0/\sqrt{n/2} = 0$ and $\operatorname{Var} Z_n = (n/2)/(n/2) = 1$.

(b)

$$\mathbb{P}[|Z_n| \geq 2] \leq \frac{\operatorname{Var} Z_n}{2^2} = \frac{1}{4}$$

(c) $1/4$ for both, since $\mathbb{P}[Z_n \geq 2] \leq \mathbb{P}[|Z_n| \geq 2]$ and $\mathbb{P}[Z_n \leq -2] \leq \mathbb{P}[|Z_n| \geq 2]$.

(d) By the Central Limit Theorem, we know that $Z_n \to \mathcal{N}(0,1)$, the standard normal distribution.

(e) Since $Z_n \to \mathcal{N}(0,1)$, we can approximate $\mathbb{P}[|Z_n| \geq 2] \approx 1 - 0.9545 = 0.0455$. By the symmetry of the normal distribution, $\mathbb{P}[Z_n \geq 2] = \mathbb{P}[Z_n \leq -2] \approx 0.0455/2 = 0.02275$.

It is interesting to note that the CLT provides a much smaller answer than Chebyshev. This is due to the fact that the CLT is applied to a particular kind of random variable, namely the (scaled) sum of a bunch of random variables. Chebyshev's inequality, however, holds for any random variable, and is therefore weaker.

3 Why Is It Gaussian?

Let X be a normally distributed random variable with mean μ and variance σ^2. Let $Y = aX + b$, where $a > 0$ and b are non-zero real numbers. Show explicitly that Y is normally distributed with mean $a\mu + b$ and variance $a^2\sigma^2$. The PDF for the Gaussian Distribution is $\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$. One approach is to start with the cumulative distribution function of Y and use it to derive the probability density function of Y.

[1.You can use without proof that the pdf for any gaussian with mean and sd is given by the formula $\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ where μ is the mean value for X and σ^2 is the variance. 2. The drivative of CDF gives PDF.]

Solution:

Problem and solution taken from *A First Course in Probability* by Sheldon Ross, 8th edition.

Let $a > 0$.

We start with the cumulative distribution function (CDF) of Y, F_Y.

$$\begin{aligned} F_Y(x) &= \mathbb{P}[Y \leq x] && \text{By definition of CDF} \\ &= \mathbb{P}[aX + b \leq x] && \text{Plug in } Y = aX + b \\ &= \mathbb{P}\left[X \leq \frac{x-b}{a}\right] && \text{Because } a > 0 \\ &= F_X\left(\frac{x-b}{a}\right) && \text{By definition of CDF. } F_X \text{ denotes the CDF of } X. \end{aligned} \tag{1}$$

Let f_Y denote the probability density function (PDF) of Y.

$$\begin{aligned} f_Y(x) &= \frac{\mathrm{d}}{\mathrm{d}x} F_Y(x) && \text{The PDF is the derivative of the CDF.} \\ &= \frac{\mathrm{d}}{\mathrm{d}x} F_X\left(\frac{x-b}{a}\right) && \text{Plug in the result from (1)} \\ &= \frac{1}{a} \cdot f_X\left(\frac{x-b}{a}\right) && \begin{array}{l}\text{PDF is the derivative of CDF.} \\ \text{Apply chain rule, } \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{x-b}{a}\right) = \frac{1}{a}.\end{array} \\ &= \frac{1}{a} \cdot \frac{1}{\sigma\sqrt{2\pi}} \cdot \mathrm{e}^{-((x-b)/a-\mu)^2/(2\sigma^2)} && X \sim \mathcal{N}(\mu, \sigma^2). \\ &= \frac{1}{a\sigma\sqrt{2\pi}} \cdot \mathrm{e}^{-(x-b-a\mu)^2/(2\sigma^2 a^2)} && \frac{x-b}{a} - \mu = \frac{1}{a}(x - b - a\mu) \end{aligned} \tag{2}$$

We have shown that f_Y equals the probability density function of a normal random variable with mean $b + a\mu$ and variance $\sigma^2 a^2$. So, Y is normally distributed with mean $b + a\mu$ and variance $\sigma^2 a^2$. The proof is done for $a > 0$. The proof for $a < 0$ is similar.