Lecture 3.2

Random variables and their distribution: Expectations

Expectation - Discrete random variables

In week 1, we summarised the sampled data using numerical values such as sample mean and sample variance.

We can also do the same for the population from which the sampled data comes, i.e., describe a random variable's distribution using numerical summaries.

One such summary it the expectation (or mean) of a random variable.

Expectation - Discrete random variables

Let X be a discrete random variable with pmf f. The **expectation** (or **expected value**) of X, denoted as $\mathbb{E}(X)$ or μ_X, is defined as

$$\mathbb{E}(X) = \sum_x x\,\mathbb{P}(X = x) = \sum_x x\,f(x).$$

Example: Roll two fair dice and let M be the largest face value showing. The pmf of M is

$$f(x) = \frac{2x-1}{36}, \quad x \in \{1, 2, \ldots, 6\}.$$

The expected value of M is

$$\mathbb{E}(M) = \sum_x x\,f(x) = 1 \times \frac{1}{36} + 2 \times \frac{3}{36} + \cdots + 6 \times \frac{11}{36} = 4.4722$$

Expectation - Continuous random variables

For continuous random variables, we can define the expectation in a similar way, replacing the pmf with pdf and sum with an integral.

Let X be a continuous random variable with pdf f . The **expectation** (or **expected value**) of X, denoted as $\mathbb{E}(X)$, is defined as

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x\, f(x)\, dx.$$

Expectation - Continuous random variables

Let X be a random variable having pdf

$$f(x) = \begin{cases} \frac{3}{4}(1 - x^2), & x \in (-1, 1) \\ 0, & x \notin (-1, 1) \end{cases}$$

The expected value of X is

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x\, f(x)\, dx = \int_{-1}^{1} x\, \frac{3}{4}(1 - x^2)\, dx = \ldots$$

Expectations of functions of random variables

Suppose X is a discrete rv with pmf f_X and g is a function. Then $Y = g(X)$ is also a discrete rv. What is $\mathbb{E}(Y)$? We need the pmf of Y, f_Y. We first recall that since $X : \Omega \to \{x_1, x_2, \ldots\}$, we have

$$\{\omega \in \Omega : g(X(\omega)) = y\} = \bigcup_{x:g(x)=y} \{\omega \in \Omega : X(\omega) = x\}\,.$$

So,

$$f_Y(y) = \mathbb{P}(Y = y) = \mathbb{P}(g(X) = y) = \sum_{x:g(x)=y} \mathbb{P}(X = x) = \sum_{x:g(x)=y} f_X(x)$$

The expectation of Y is then

$$\begin{aligned}\mathbb{E}Y &= \sum_y y\, f_Y(y) = \sum_y y \sum_{x:g(x)=y} f_X(x) = \sum_y \sum_{x:g(x)=y} y\, f_X(x) \\ &= \sum_x g(x)\, f_X(x)\,.\end{aligned}$$

Expectations of functions of random variables

These previous calculations, lead to the following important result:

Law of the Unconscious Statistician (LOTUS)

- If X is discrete with pmf f, then for any real-valued function g

$$\mathbb{E}(g(X)) = \sum_{x} g(x)\, f(x).$$

- If X is continuous with pdf f, then for any real-valued function g

$$\mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(x)\, f(x)\, dx.$$

Linearity of expectations

The expectation is "linear", i.e., for any functions g and h, and any real numbers a and b

$$\mathbb{E}(a \cdot g(X) + b \cdot h(X)) = a \cdot \mathbb{E}(g(X)) + b \cdot \mathbb{E}(h(X)).$$

For example, $\mathbb{E}(aX + b) = a\mathbb{E}(X) + b$.

Example: Suppose the rv X models the temperature, in $^\circ F$, of a CPU with $\mathbb{E}X = 130$. What is the expected temperature in $^\circ C$?

Answer: We know that $Y = (X - 32)/1.8$ models the temperature in $^\circ C$ of the CPU.

$$\mathbb{E}Y = \mathbb{E}\left(\frac{X - 32}{1.8}\right) = \frac{\mathbb{E}X}{1.8} - \frac{32}{1.8} = 54.44$$

Variance

Another useful numerical characteristic of the distribution of X is the *variance* of X. This number, sometimes written as σ_X^2, measures the *spread* or dispersion of the distribution of X.

The **variance** of a random variable X is

$$\mathsf{Var}(X) = \mathbb{E}(X-\mu)^2 \ ,$$

where $\mu = \mathbb{E}(X)$. The square root of the variance is called the **standard deviation**.

For any random variable X

$$\begin{aligned} \mathsf{Var}(X) = \mathbb{E}(X-\mu)^2 &= \mathbb{E}(X^2 - 2\mu X + \mu^2) \\ &= \mathbb{E}X^2 - 2\mu\mathbb{E}X + \mu^2 = \mathbb{E}X^2 - \mu^2 \end{aligned}$$

Nonlinearity of Variance

Recall that expectation is linear, i.e., any real numbers a and b, $\mathbb{E}(a\,X + b) = a\,\mathbb{E}(X) + b$. However, the variance is "nonlinear", i.e., any real numbers a and b

$$\begin{aligned}\mathrm{Var}(a\,X + b) &= \mathbb{E}\left((a\,X + b) - (a\,\mu + b)\right)^2 \\ &= \mathbb{E}\left(a^2(X-\mu)^2\right) = a^2\,\mathrm{Var}(X)\end{aligned}$$

Example: Suppose the rv X models the temperature, in $^\circ F$, of a CPU with $\mathrm{Var}(X) = 80$. What is the variance of temperature in $^\circ C$?

Answer: We know that $Y = (X - 32)/1.8$ models the temperature in $^\circ C$ of the CPU.

$$\mathrm{Var}\,Y = \mathrm{Var}\left(\frac{X-32}{1.8}\right) = \frac{\mathrm{Var}X}{(1.8)^2} = \frac{80}{(1.8)^2} = 24.69.$$