Math 170E Week 9

Aidan Jan

March 13, 2023

Distribution Functions

$F_X : \mathbb{R} \to \mathbb{R}$

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

Note that F_X is always (1) nondecreasing, and (2) $F_X(+\infty) = 1$, $F_X(-\infty = 0)$.
Example:
Assume that $X \sim \text{unif}[0, 2]$, and has the density function:

$$F_X(a) = \begin{cases} 1 & a > 2 \\ \frac{a}{2} & 0 \leq a \leq 2 \\ 0 & a < 0 \end{cases}$$

The graph of F_X would look like the following:
INSERT IMAGE HERE
Example:
Assume $X \sim \text{Bern}(0.4)$

$$F_X(a) = \mathbb{P}(X \leq a) = \begin{cases} 0 & a < 0 \\ 0.6 & a = 0 \\ 0 & 0 < a < 1 \\ 1 & a \geq 1 \end{cases}$$

INSERT IMAGE HERE
Example:
Let $w \in \{1, 2, 3\}$. Then, $w < 1.5 \Leftrightarrow w = 1$. Also, $w < 1.999 \Leftrightarrow w = 1$. This results in a step function graph like the previous example.

Why do we need distribution functions?

1. They are well-defined for every random variable. Whether they are discrete or continuous random variables, they make sense for both types. This is unlike probability mass functions, which do not make sense for continuous random variables.
2. Easy to define concepts. It is easy to use distribution functions to define mathematical representations

of scenarios.
3. Not easy for calculation. A lot of times, it is necessary to know the probability of landing on or below certain numbers. Some functions defining the distribution are not easy to work with.

Distribution Functions of Multiple Random Variables

Consider $F_{X,Y}(a,b) = \mathbb{P}(X \leq a \,\&\, Y \leq b)$. The following is also true if X and Y are independent:

$$X \perp Y \Leftrightarrow \forall\, a, b\, F_{X,Y}(a,b) = F_X(a) \cdot F_Y(b)$$

The other rules of independence also work, such as $\{X \cdots\} \perp \{Y \cdots\}$ and $f(X) \perp g(Y)$.

Probability Density Functions

Example:
Consider gravity. Your weight on Earth is determined by the equation $\frac{M \cdot m}{r^2}$, where M is the mass of Earth, m is your mass, and r is the radius of the Earth. But how does this work?
If you divide the Earth into small chunks, you will find that each chunk is not the same distance to you. In fact, the distances can vary by a lot. Therefore, the total gravity can be written as the summation of the force of gravity on you from each individual chunk. Additionally, the Earth is not homogeneous; its density varies depending on where you are: the Crust has the lowest density while the Core has the highest. The density on something can be written as its weight divided by its volume, thus total gravity can be written as an integral over all the tiny chunks of Earth using the density function for all the chunks.

Consider finding the density of a point, a. Let X be a random variable which picks the location of the point, and f represent the density function. Thus, $f_X(a)$ tells the density at a given point, a. However, something with no volume cannot have a density, so we must use a region, and set a limit to make the region as close as possible to the point.
Let ϵ represent an area around the point. Thus,

$$f_X(a) = \lim_{\epsilon \to 0} \frac{\mathbb{P}(X \in [a-\epsilon, a+\epsilon])}{|[a-\epsilon, a+\epsilon]|}$$

This is essentially a derivative. By definition, the cdf. of X can be written as

$$F_X(a) = \frac{\mathbb{P}(X \leq a+\epsilon) - \mathbb{P}(X \leq a-\epsilon)}{2\epsilon}$$

If the derivative of this is taken, we get:

$$F'_X(a) = \lim_{\epsilon \to 0} \frac{F_X(a+\epsilon) - F_X(a-\epsilon)}{2\epsilon}$$

This is not a rigorous proof, but for this class, we can assume that $f_X(a) := F'_X(a)$. Note that they are not completely equal, they would be close to each other, other than a few points.

What if $X \sim \text{Bern}(p)$? In this case, the cdf. of X is not continuous. Thus, X has no f_X.
If $X \sim$ unif, then we have a continuous, piecewise function for F_X. In this case, there is a derivative, F'_X

will be piecewise continuous.

Example:
Suppose $X \sim \text{unif}[0,1]$. What is $f_X(0)$?
In this case,

$$f_X(a) = \begin{cases} 0 & a \notin [0,1] \\ 1 & a \in (0,1) \end{cases}$$

Example:
Suppose that

$$f_X = \begin{cases} c \cdot \sqrt{4 - X^2} & |X| \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

What is c? First, we should observe some properties of density functions:
(1). $f \leq 0$
(2). $\int_{-\infty}^{+\infty} f_X(a)\, \mathrm{d}a = 1$
(3). $\int_{-\infty}^{+\infty} F_X'(a)\, \mathrm{d}a = F_X(a)|_{-\infty}^{+\infty} = 1$

For this problem, we know that the integral of $f_X(a)$ must equal to 1. Therefore,

$$\int_{-2}^{2} c \cdot \sqrt{4 - a^2}\, \mathrm{d}a = 2\pi c = 1$$

(The $\sqrt{4 - a^2}$ is the equation of a semicircle, radius 2) Therefore, $c = \frac{1}{2\pi}$.

Lemma:
For X, if f_X exists and piecewise continuous, then

$$\mathbb{P}(X \in [a,b]) = \int_a^b f_X(s)\, \mathrm{d}s = \int_a^b F_X'(s)\, \mathrm{d}s = F(b) - F(a)$$

Gaussian Random Variables

This is a random variable with a Gaussian distribution, a.k.a., a normal distribution, or a bell curve.
Example:
$X \sim N(\mu, \sigma^2)$
Then, by definition,

$$f_X(a) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(a-\mu)^2}{2\sigma^2}}$$

The expected value and variance are

$$\mathbb{E}X = \mu \qquad \text{Var}X = \sigma^2$$

X is defined everywhere on $\mathbb{R}$ and is symmetric over line $x = \mu$. The width of the "peak" section is approximately 2σ.
INSERT IMAGE HERE
The Gaussian distribution has some properties that makes it unique from other distributions. These include:

1. $X \sim N(\mu, \sigma^2)$ is linear. This means that $\forall a, b$, $aX + b$ is also a normal distribution.
2. If X, Y have normal distributions, and $X \perp Y$, then $X + Y$ also has a normal distribution.

Example:
Suppose that $X \sim N(240, 100)$. What is $\mathbb{P}(X > 250)$?

To solve this, let $Y = \frac{X-240}{10}$. In this case, $\mathbb{P}(X > 250) = \mathbb{P}(Y > 1)$. Since X is a Gaussian random variable, Y is also a Gaussian random variable. In this case, $Y \sim N(0, 1)$, which is the **standard normal distribution**. From here, you use a calculator.

Example:
Let $X \sim N(70, 5)$ describe the distribution of heights for average adult males and $Y \sim N(65, 4)$ describe the distribution of heights for average adult females. What is $\mathbb{P}(|X - Y| \geq 6)$?

Since Gaussian distributions are linear, $|X - Y| \sim N(5, \sqrt{41})$. (Remember to add standard deviations, you sum the squares.)

Example:
Reconsider the Newborn problem. An average of k babies are born in a hospital every day, with an approximation given by a Poisson(k). Let $X \approx$ Poisson(3). $\mathbb{P}(X = k) \approx \mathbb{P}(Y = k), Y \sim N(\dots)$.

$$\mathbb{P}(X \in [a, b]) \approx \mathbb{P}(Y \in [a, b]), \forall a, b$$

It turns out that Poisson distribution curves and Gaussian distribution curves are very similar. The difference is that Poisson is for discrete random variables, while Gaussian is for continuous random variables. They can both be written as $X = \sum_{k=1}^{N} X_k$, where in Gaussian, X_k can be any real number, while in Poisson, X_k must be an integer.