Stats 130
Day 8 Notes

Elijah Hantman

Poisson

A Poisson random variable counts the number of occurances of an event. It has no upper bound.

$$f_X(x) = \frac{e^{-\lambda}\lambda^x}{x!}, x = 0, 1, 2, ...$$

Here λ is the is the intensity of the process and is proportional to the number of expected occurances in a given time frame.

$f_X(x)$ is greater than 0 by defninition. $\forall x$

$$\sum_{x=0}^{\infty} f_X(x) = \sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} = e^{-\lambda}e^{\lambda} = 1$$

This shows the Poisson is a probabiilty function, and is denoted $X \sim Pois(\lambda)$

Example: The number of customers that order a pizza in a 30 min interval is a Poisson random variable with $\lambda = 5$. What is the probability that between 6 and 6:30 three or more customers will order pizza?

$$f_X(x) = \frac{e^{-\lambda}\lambda^x}{x!} \tag{1}$$

$$f_X(x) = \frac{e^{-5}5^x}{x!}, x = 0, 1, 2 \tag{2}$$

$$Pr(X \geq 3) = 1 - (Pr(X = 0) + Pr(X = 1) + Pr(X = 2)) \tag{3}$$

$$= 1 - e^{-5}(1 + 5, + \frac{25}{2}) = 0.875348 \tag{4}$$

Continuous Random Variables

Reimagine the building power and water example. If we no longer require the demand to be a discrete integer value we now have a continuous random variable. There are an infinite number of elements in the Sample Space, and they cannot be traversed in a countable order.

Cumulative Distribution Function

With continuous random variables, the idea of $X = x$ is unhelpful, since $Pr(X = x)|X : S \to D \subseteq \mathbb{R}$ is always equal to zero or some infinitesimal. For continuous values it is more helpful and tractable to work with ranges.

A CDF is defined as

$$F_X(x) = Pr(X \leq x)$$

A CDF can be made for any random variable whether continuous or discrete. The capital letters are also part of the notation.

Some properties of CDFs are:

- $0 \leq F_X(x) \leq 1, \forall x$
- $F_X(x)$ is non decreasing.
  $x_1 < x_2 \implies X \leq x_1 \subset X \leq x_2$ therefore, $Pr(X \leq x_1) \leq Pr(X \leq x_2)$ which implies $F_X(x_1) \leq F_X(x_2)$
- $\lim_{x \to -\infty} F_X(x) = 0, \lim_{x \to +\infty} F_X(x) = 1$
- For $b > a$, $Pr(X \in (a, b]) = Pr(a < X \leq b)$
  $= Pr(X \leq b) - Pr(X \leq a) = F_X(b) - F_X(a)$

If $F_X(x)$ is continuous for all x, then X is a continuous random variable.

Imagine a discrete random variable with probability function and CDF as follows.

$$f_X(x) = \begin{cases} \frac{1}{k} & x = 1, 2, 3..., k \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

$$F_X(x) = Pr(X \leq x) = \sum_{i \leq x} Pr(X = i) = \begin{cases} 0 & x < 1 \\ \frac{\lceil x \rceil}{k} & 1 \leq x < k \\ 1 & x \leq k \end{cases} \tag{6}$$

This CDF is not continuous which implies the random variable must be discrete.

Probability Density Function

When X is a continuous random variable, $F_X(x)$ is both continuous and differentiable. Therefore we can calculate the derivative of $F_X(x)$.

The Probability Density Function or PDF, is defined as,

$$f_X(x) = F'_X(x)$$

A PDF is not defined for discrete random variables since the CDF is not continuous or differentiable.

In addition the lower case "f" is used since the discrete meaning of a probability function is undefined for continuous variables. This means only one possible meaning is valid at a time which is unambiguous.

Properties:

- $$F_X(x) = \int_{-\infty}^{x} f_X(t)dt \quad \forall x$$
- $$Pr(a < x \leq b) = F_X(b) - F_X(a) \tag{7}$$
  $$= \int_{-\infty}^{b} f_X(t)dt - \int_{-\infty}^{a} f_X(t)dt \tag{8}$$
  $$= \int_{a}^{b} f_X(t)dt \tag{9}$$
- $$f_X(x) \geq 0, \forall x$$
- $$Pr(X \in A) = \int_A f_X(t)dt$$
- $$\int_{-\infty}^{\infty} f_X(t)dt = \lim_{x \to +\infty} F_X(x) = 1$$

Examples

Utility problem. Let X be the water demand and suppose the probability of an event is proportional to the area of the event.

$$F_X(x) = Pr(X \leq x)\frac{(150-1)\times(x-4)}{(150-1)\times(200-4)} = \frac{x-4}{196}$$

The bottom is the area of the sample space, the top is the area that represents the interval we are interested in.

The pdf of this function is

$$f_X(x) = \begin{cases} \frac{1}{196} & 4 \leq x \leq 200 \\ 0 & \text{otherwise} \end{cases}$$

The pdf is constant, which is a distribution known as a uniform distribution. A uniform distribution is notated $X \sim Unif(a, b)$

Consider a random variable X with pdf

$$f_X(x) = \begin{cases} cx & 0 < x < 4 \\ 0 & \text{otherwise} \end{cases}, c > 0$$

Determine the value of c. Calculate $Pr(1 \leq X \leq 2)$. The integral of $f_X(x)$ must equal 1.

$$\int_{-\infty}^{+\infty} f_X(t)dt = \int_0^4 ctdt = \left.\frac{ct^2}{2}\right|_0^4 = 8c$$

Therefore $c = \frac{1}{8}$

$$Pr(1 \leq X \leq 2) = \frac{3}{16}$$