Poisson Distribution

We say that a random variable X has a Poisson distribution when X can take on any non-negative integer, and the probability mass function is given by

$$P(X=k)=e^{-\lambda}\frac{\lambda^k}{k!},\quad k=0,1,2,\ldots\quad \text{for some } \lambda>0.$$

We have that $EX=\lambda$ and $\operatorname{var}X=\lambda$.

Example 1. Suppose that the number of visitors who visit a certain website is Poisson distributed with mean 20 visitors/minute.

(a) What is the probability that in the next fifteen minutes there will be exactly 250 visitors?

(b) What is the probability that in the next five minutes nobody will visit the website?

(c) What is the probability that there will be 90 visitors in the next five minutes and then 160 visitors in the ten minutes after that?

(d) What is the expected number of visitors in the next hour? What is the variance?

Solution.

(a) The number of visitors in the next fifteen minutes is Poisson distributed with mean $20\times 15=300$. So the probability that there are 250 visitors is $e^{-300}\frac{300^{250}}{250!}$.

(b) The number of visitors in the next five minutes is Poisson distributed with mean $20\times 5=100$. So the probability that there is 0 visitor is $e^{-100}\frac{100^0}{0!}=e^{-100}$. (Recall that $0!=1$.)

(c) Let X be the number of visitors in the next five minutes, and Y the number of visitors in the ten minutes after that. Then X and Y are Poisson distributed. Moreover, they are independent since the time periods do not overlap. So

$$\begin{aligned}P(X=90,Y=160)&=P(X=90)P(Y=160)\\&=\left(e^{-100}\frac{100^{90}}{90!}\right)\left(e^{-200}\frac{200^{160}}{160!}\right).\end{aligned}$$

(d) The expected number is simply $20\times 60=1200$. Since the number of visitors is Poisson distributed, the variance is the same as the expected number, so the variance is also 1200. ■

Remark. If X and Y are two independent Poisson variables with mean λ_X and λ_Y, respectively, then $X+Y$ is also Poisson with mean $\lambda_X+\lambda_Y$.

Poisson Approximation. When X is *binomial* with large n and small p,

$$P(X=k)\approx e^{-\lambda}\frac{\lambda^k}{k!},\quad \lambda=np.$$

That is, we think of X as a Poisson variable with $\lambda=np$ to approximate $P(X=k)$.

Example 2. About 1 in 900 men have Klinefelter syndrome. Among the randomly selected 1,800 men, what is the probability that at least two of them have this syndrome?

Solution. Let X be the number of people who have Klinefelter syndrome among this group of 1,800 men. Then X is binomial with $n=1800$ and $p=1/900$. We want to find $P(X\geq 2)$, so

$$\begin{aligned}P(X\geq 2)&=1-P(X=0)-P(X=1)\\&=1-\left(\frac{899}{900}\right)^{1800}-\binom{1800}{1}\left(\frac{1}{900}\right)\left(\frac{899}{900}\right)^{1799}.\end{aligned}$$

Since n is large and p is small, we can also use Poisson approximation with $\lambda=np=1800\cdot\frac{1}{900}=2$:

$$\begin{aligned}P(X\geq 2)&=1-P(X=0)-P(X=1)\\&\approx 1-e^{-2}\frac{2^0}{0!}-e^{-2}\frac{2^1}{1!}=1-3e^{-2}.\end{aligned}$$ ■

Comparison

distribution	parameters	values	$P(X=k)$	EX	$\operatorname{var}X$
binomial	n,p	$0,1,\ldots,n$	$\binom{n}{k}p^k(1-p)^{n-k}$	np	$np(1-p)$
geometric	p	$1,2,\ldots$	$(1-p)^{k-1}p$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
Poisson	λ	$0,1,2,\ldots$	$e^{-\lambda}\frac{\lambda^k}{k!}$	λ	λ

Continuous Random Variables - Basics

X is a continuous random variable when it takes on a continuum of values. It is characterized by a *(cumulative) distribution function* $F(x)$.

$$F(x) = P(X \leq x).$$

The *(probability) density function* is the function $f(x)$ such that

$$F(x) = \int_{-\infty}^{x} f(u)du.$$

To find the density function for a continuous random variable, simply take the derivative of the distribution function:

$$f(x) = F'(x).$$

(For x where F is not differentiable, simply set $f(x) = 0$.)

A function is a valid density function if **(i)** it is non-negative (i.e. $f(x) \geq 0$ for all x) and **(ii)** $\int_{-\infty}^{\infty} f(u)du = 1$.

We can use the density function to find the probability of the random variable over an interval:

$$P(a \leq X \leq b) = \int_{a}^{b} f(u)du.$$

We use the density function to find the mean and the variance:

$$EX = \int_{-\infty}^{\infty} uf(u)du. \qquad E(X^2) = \int_{-\infty}^{\infty} u^2 f(u)du.$$
$$\operatorname{var} X = E(X^2) - (EX)^2.$$

Example 3. Suppose that X is a continuous random variable with the following distribution function:

$$F(x) = \begin{cases} 0 & x < 0 \\ 2x - x^2 & 0 \leq x \leq 1 \\ 1 & x > 1. \end{cases}$$

Compute (a) $P(X \leq 0.3)$, (b) $P(X > 0.7)$, (c) $P(-0.2 < X \leq 1.4)$, (d) EX, and (e) $\operatorname{var} X$.

Solution.

(a) $P(X \leq 0.3) = F(0.3) = 2(0.3) - (0.3)^2 = 0.51.$

(b) $P(X > 0.7) = 1 - P(X \leq 0.7) = 1 - F(0.7) = 1 - [2(0.7) - (0.7)^2] = 0.09.$

(c) $P(-0.2 < X \leq 1.4) = P(X \leq 1.4) - P(X \leq -0.2) = F(1.4) - F(-0.2) = 1 - 0 = 1.$

(d) To find EX and $\operatorname{var} X$, we must find the density function first.

$$f(x) = F'(x) = \begin{cases} 2 - 2x & 0 \leq x \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

$$EX = \int_{-\infty}^{\infty} uf(u)du = \int_{0}^{1} u(2-2u)du = \int_{0}^{1} 2u - 2u^2 du = \left[u^2 - \frac{2u^3}{3}\right]_0^1 = \frac{1}{3}.$$

(e) To find the variance, we first find $E(X^2)$.

$$\begin{aligned} E(X^2) &= \int_{-\infty}^{\infty} u^2 f(u)du \\ &= \int_{0}^{1} u^2(2 - 2u)du = \int_{0}^{1} 2u^2 - 2u^3 du = \left[\frac{2u^3}{3} - \frac{u^4}{2}\right]_0^1 = \frac{2}{3} - \frac{1}{2} = \frac{1}{6}. \end{aligned}$$

Then $\operatorname{var} X = E(X^2) - (EX)^2 = \frac{1}{6} - \left(\frac{1}{3}\right)^2 = \frac{1}{18}$. ■

Example 4. Suppose that X is a continuous random variable with the following density function:

$$f(x) = \begin{cases} 0.5 & 0 \leq x < 0.5 \\ c & 0.5 \leq x < 1 \\ 0 & \text{otherwise} \end{cases}$$

where c is a real number. Find the value of c, then compute $P(0.3 < X < 0.8)$.

Solution. For a function f to be a density function, it must be non-negative and $\int_{-\infty}^{\infty} f(u)du = 1$. So we know that c must be non-negative, and that

$$1 = \int_{-\infty}^{\infty} f(u)du = \int_{0}^{0.5} 0.5du + \int_{0.5}^{1} cdu = 0.5(0.5) + 0.5c = 0.25 + 0.5c.$$

So $c = 0.75/0.5 = 1.5$. Then

$$\begin{aligned} P(0.3 < X < 0.8) &= \int_{0.3}^{0.8} f(u)du = \int_{0.3}^{0.5} 0.5du + \int_{0.5}^{0.8} 1.5du \\ &= (0.5 - 0.3)0.5 + (0.8 - 0.5)1.5 = 0.55. \quad ■ \end{aligned}$$