# Poisson Random variables and the connection to the Binomial distribution.

Poisson random variable arises out of an "approximation" to the Binomial distribution, as the number of trials $n \to \infty$.

Let success prob $= p$, $np = \lambda$ (some constant $>0$) for a Binomial process.

$\hookrightarrow p = \frac{\lambda}{n}$

$$P(X=i) = \binom{n}{i} p^i (1-p)^{n-i} =$$

$$= \frac{n!}{(n-i)!\, i!} \left(\frac{\lambda}{n}\right)^i \left(1 - \frac{\lambda}{n}\right)^{n-i}$$

$$= \underbrace{\frac{n(n-1)(n-2)\cdots(n-i+1)}{n^i}}_{\to 1} \frac{\lambda^i}{i!} \underbrace{\frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^i}}_{\to e^{-\lambda}/1}$$

$$\to \frac{e^{-\lambda}\lambda^i}{i!} \quad \text{as } n \to \infty \quad \left(\text{we have used } \lim_{n\to\infty} \left(1-\frac{\lambda}{n}\right)^n = e^{-\lambda}\right)$$

Next, note that Poisson's approximation is by itself a probability function:

$$P(X=x) = \frac{e^{-\lambda}\lambda^x}{x!} = f_p(x) \quad ; \quad 0 \le f_p(x) \le 1$$

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

Def | A discrete random variable $X$ with (countably infinite) values $0, 1, 2, \ldots$ is called Poisson with parameter $\lambda > 0$ if:

$$P(X=x) = \frac{e^{-\lambda}\lambda^x}{x!} = f_\lambda^P(x) \quad ; \quad x = 0, 1, 2, \ldots$$

The expected value is:

$$E[X] = \sum_{x=0}^{\infty} x P(X=x) = \sum_{x=1}^{\infty} x \frac{e^{-\lambda}\lambda^x}{x!}$$

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

$$= \lambda e^{-\lambda}e^{\lambda} = \lambda \quad \Rightarrow \boxed{E[X] = \lambda}$$

$\boxed{Var[X] = \sigma_X^2 = \lambda}$ as well! $\left(\sigma_X^2 = E[(X-\lambda)^2]\right)$

Ex) Suppose that flaws in some wooden pieces used for furniture occur at random with an avg of 1 flaw per $50\,cm^2$. What is the probability that a $4\,cm \times 8\,cm$ small sheet has (a) no flaws, (b) at most one flaw?

Assume the # of flaws per unit area is Poisson distributed. Let $\underline{X \text{ be Poisson r.v.}}$ recording # of flaws in $4 \times 8\,cm^2$ sheet.

What is $\lambda$? Recall $\lambda = E[X]$.

We have $4 \times 8\,cm^2 = 32\,cm^2$ sheet
We expect one flaw per $50\,cm^2$.

$$\Rightarrow E[X] = \frac{32}{50} \text{ flaws} < 1 \text{ flaw}$$

$$\text{Let } \lambda = \frac{32}{50} = \frac{16}{25}$$

$$P(X=x) = \frac{e^{-\lambda}\lambda^x}{x!}$$

$$P(X=0) = \frac{e^{-\frac{16}{25}}\left(\frac{16}{25}\right)^0}{0!} = e^{-\frac{16}{25}} \approx 0.53$$

$P(\text{at most one flaw}) = P(X \le 1) = P(X=0) + P(X=1)$

$X=0, X=1$ are mutually exclusive events

$$P(X \le 1) = e^{-16/25} + \frac{e^{\frac{-16}{25}}\left(\frac{16}{25}\right)^1}{1!}$$

$$= e^{-0.64} + 0.64e^{-0.64} \approx 0.86$$

Ex) Suppose that on avg, in every 3 pages of a book there is one typo. If the # of typos on a single page of a book is a Poisson r.v., what is the probability of at least one error on a certain page of a book?

$X$ r.v. (Poisson) counting # of errors on the page we are interested in.

$$E[X] = \frac{1}{3} = \lambda \quad ; \quad P(X=x) = \frac{\left(\frac{1}{3}\right)^x e^{-1/3}}{x!}$$

$$P(X \ge 1) = 1 - P(X=0) = 1 - e^{-1/3} \approx 0.28$$

# Geometric random variable

Suppose that a sequence of independent Bernoulli trials (Binomial trials), each with probability of success $0 < p < 1$, are performed.

Let $X$ be the # of trials until the first success occurs. Then $X$ is a discrete random variable called <u>geometric</u>. It's set of possible values is $1, 2, 3, \ldots$ (countably infinite) and

$$P(X=x) = (1-p)^{x-1} p = f_p^g(x)$$

[ note, no limit on # of trials prior to first success

$(x-1)$ trials are failures, $x$-th trial is a success, successive Bernoulli trials are independent.

$$\sum_{x=1}^{\infty} p(x) = \sum_{x=1}^{\infty} (1-p)^{x-1} p = \frac{p}{1-(1-p)} = 1$$

since $\sum_{x=1}^{\infty} (1-p)^{x-1}$ is a geometric series.

Hence, since $0 \le p(x) \le 1$, $p(x)$ is a probability function.

$$E[X] = \sum_{x=1}^{\infty} x f_p^g(x) = \sum_{x=1}^{\infty} x p (1-p)^{x-1} =$$

$$= \frac{p}{1-p} \sum_{x=1}^{\infty} x (1-p)^x = \frac{p}{1-p} \frac{1-p}{[1-(1-p)]} = \frac{1}{p}$$

where we have used $\sum_{x=1}^{\infty} x r^x = \frac{r}{(1-r)^2}$, $|r|<1$.

Can show that:

$$Var[X] = E[X^2] - E[X]^2 = \frac{2-p}{p^2} - \left(\frac{1}{p}\right)^2$$

$$= \frac{1-p}{p^2} = \frac{1}{p^2} - \frac{1}{p}$$

where $E[X^2] = \sum_{x=1}^{\infty} x^2 p (1-p)^{x-1}$

Ex) From a deck of 52 cards, we draw cards at random until ace is drawn. What is the probability at least 10 draws are needed? Let X be a geometric r.v. We must then "extend" the problem to allow arbitrary many trials, see below.

$$E[X] = \frac{1}{p} = 13 \left(= \frac{4}{52}\right) \Rightarrow p = \frac{1}{13}$$

Thus,

$$P(X=x) = \left(\frac{12}{13}\right)^{x-1}\left(\frac{1}{13}\right), \quad x=1,2,3,\ldots$$

$$P(X\geq 10) = \sum_{n=10}^{\infty}\left(\frac{12}{13}\right)^{n-1}\left(\frac{1}{13}\right) = \left(\frac{12}{13}\right)^{9} \approx 0.49$$

same as probability of no aces in first nine <u>blind</u> draws.

Note that the maximum number of draws needed is at most 48.

$$\sum_{n=10}^{\infty}\left(\frac{12}{13}\right)^{n-1}\left(\frac{1}{13}\right) = \frac{1}{13}\sum_{n=10}^{\infty}\left(\frac{12}{13}\right)^{n-1}$$

$$= \frac{1}{13}\;\frac{(12/13)^9}{1-12/13} = \left(\frac{12}{13}\right)^{9} \approx 0.49$$

Geometric r.v. can be used assuming we combine many decks together and draw cards from an infinite pool.

Recall <u>geometric series</u>:

$$\sum_{n=0}^{\infty} ar^n = a + ar + ar^2 + \ldots \quad \text{converges for } |r|<1.$$

When $|r|<1$, the sum of $\infty$ many terms is $\dfrac{a}{1-r}$

$$S_n = a + ar + \ldots + ar^{n-1} \quad \text{(partial sum, n terms)}$$

$$rS_n = ar + ar^2 + \ldots + ar^n$$

$$S_n - rS_n = a - ar^n$$

$$\Rightarrow S_n(1-r) = a(1-r^n) \Rightarrow S_n = a\left(\frac{1-r^n}{1-r}\right)$$

$$S_n = \frac{a}{1-r} - \frac{ar^n}{1-r}$$

For $n \to \infty$ and $|r| < 1 \Rightarrow S_n \to \frac{a}{1-r}$

Thus, e.g.:

$$\sum_{n=10}^{\infty} \left(\frac{12}{13}\right)^{n-1} = \frac{(12/13)^9 \leftarrow \text{first term } (a)}{1 - 12/13 \leftarrow r}$$

In order for the event $X = x$ to occur, it is necessary to have $x-1$ failures followed by a success:

$$f_p^g(x) = P(X_g = x) = p\,q^{x-1} \quad \text{where } q = 1-p$$

$$\sum_{x=1}^{\infty} f_p^g(x) = p \sum_{x=1}^{\infty} q^{x-1} = p(1 + q + q^2 + \dots)$$
$$= p\,\frac{1}{1-q} = \frac{p}{p} = 1$$

## No memory property

Suppose $X \sim Geo(p)$, then:

$$P[X > j+k \mid X > j] = P[X > k]$$