Probability Distribution Functions - Discrete variable

Josep L. Carrasco
Bioestadística. Departament de Fonaments Clínics
Universitat de Barcelona

Discrete random variable: functions and properties

- Variable: counts of events.
- Example: number of women in 2nd year of the Biomedical Engineering degree.

Probability mass function

The probability mass function assigns probability to each possible value of the random variable.

$$\begin{aligned} P : \mathbb{R} &\longrightarrow [0,1] \\ k &\longrightarrow P\left(X = k\right) = p \end{aligned}$$

$$P\left(X = k\right) \in [0,1]$$
$$\sum_{i=-\infty}^{+\infty} P\left(X = k\right) = 1$$

Example: number of boys in sets of 3 siblings. Assuming equiprobability and independence among births:

X = number of boys.

$P\left(\female\right) = P\left(\male\right) = 0.5$

Probability Distribution Function

The probability distribution function assigns **cumulative probability** to each possible value of the random variable.

Ω	X	p
♀♀♀	0	0.125
♀♀♂	1	0.125
♀♂♀	1	0.125
♂♀♀	1	0.125
♀♂♂	2	0.125
♂♀♂	2	0.125
♂♂♀	2	0.125
♂♂♂	3	0.125

X	$P(X=k)$
0	0.125
1	0.375
2	0.375
3	0.125

$$\begin{aligned} F:\mathbb{R} &\longrightarrow [0,1] \\ k &\longrightarrow F(k) = P(X \leq k) = p \end{aligned}$$

$$P(k_1 < X \leq k_2) = F(k_2) - F(k_1)$$

k	$P(X=k)$	$P(X \leq k)$
0	0.125	0.125
1	0.375	0.5
2	0.375	0.875
3	0.125	1

Expectation of a Random Variable $E(X)$

Expectation: the expected value of a random variable.

Prediction: a guess about the value of the random variable.

Prediction error: the difference between the prediction and the actual value of the random variable.

The expectation **minimizes the prediction error** when the number of predictions tend to infinite.

Discrete variable: expectation appears as a weighted mean.

$$E(X) = \sum_{k=-\infty}^{+\infty} k \cdot P(X = k)$$

- If we randomly selected a set of 3 siblings, how many boys do we expect to find?

$$E(X) = \sum_{k=0}^{3} k \cdot P(X = k) =$$

$$0 \cdot P(X = 0) + 1 \cdot P(X = 1) + 2 \cdot P(X = 2) + 3 \cdot P(X = 3) =$$

$$0 \cdot 0.125 + 1 \cdot 0.375 + 2 \cdot 0.375 + 3 \cdot 0.125 = 1.5$$

The expectation is 1.5 boys. Notice that the expectation may be a value that the random variable cannot take.

Properties of the Expectation

Let X and Y be two random variables and C a constant.

- Linearity.

$$E(X + C) = E(X) + C$$
$$E(X + Y) = E(X) + E(Y)$$

- Multiplicity.

$$E(X \cdot C) = E(X) \cdot C$$

- If X and Y are independent:

$E(X \cdot Y) = E(X) \cdot E(Y)$

Variance of a random variable

- Variance: indicates the dispersion of the variable around its expectation.

It is a statistical distance:

- How far are the values from the expectation?
- What is the mean distance of the random variable to its expectation?
- How much representative of the random variable is the expectation?

If the **variance is large** the data will be very **spread** and the expectation can not be taken as representative of the random variable.

- General expression

$$V\left(X\right)=E\left[\left(X-E\left(X\right)\right)^2\right]=E\left(X^2\right)-E\left(X\right)^2$$

- Variance in discrete variables

$$V\left(X\right)=\sum_{k=-\infty}^{+\infty}k^2\cdot P\left(X=k\right)-E\left(X\right)^2$$

- Example: number of boys in 3 siblings

k^2	$P\left(x=k\right)$	$k^2 P\left(x=k\right)$
0	0.125	0
1	0.375	0.375
4	0.375	1.5
9	0.125	1.125
		$\sum=3$

$$V\left(X\right)=\sum_{k=0}^{3}k^2 P\left(X=k\right)-E\left(X\right)^2=3-1.5^2=0.75$$

The random variable "number of boys in 3 siblings" has a expectation of 1.5 boys and a variance of 0.75 boys^2.

Properties of the Variance

Let X and Y be two random variables and C a constant.

- Invariant to changes in location parameter

$$V\left(X+C\right)=V\left(X\right)$$

- Scale

$$V\left(X\cdot C\right)=C^2\cdot V\left(X\right)$$

- Variance of the sum of two random variables

$$V(X+Y) = V(X) + V(Y) + 2Cov(X,Y)$$

where $Cov(X,Y)$ is the covariance which is a measure of *joint variability*.

$$Cov(X,Y) = E(XY) - E(X)E(Y)$$

If two variables, X and Y, are *independent*, $Cov(X,Y) = 0$.
So that, if two variables, X and Y, are independent:

$$V(X+Y) = V(X) + V(Y)$$

- Variance of the difference of two random variables

$$V(X-Y) = V(X) + V(Y) - 2Cov(X,Y)$$

Probability distribution models

Bernoulli

- Binary results. The variable usually takes values 0 or 1.
- Probability distribution defined by one single parameter $p = P(X = 1)$.
- Probability mass function

$$\begin{aligned} P(X = 1) = & \; p \\ P(X = 0) = & \; 1 - p \end{aligned}$$

- Expectation and Variance

$$\begin{aligned} E(X) = & \; p \\ V(X) = & \; p(1 - p) \end{aligned}$$

Example. Flipping a coin. X= "Head side"

$$p = 0.5, \; E(X) = 0.5, \; V(X) = 0.5 \cdot 0.5 = 0.25$$

Binomial

The addition of n independent and identically distributed Bernoulli variables.

$$X = \sum_{i=1}^{n} Y_i, \qquad Y_i \sim \text{Bernoulli}(p)$$

- Range of values: $\{0, 1, \ldots, n\}$
- Probability mass function

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

- Expectation and Variance

$$\begin{aligned} E(X) = & \; np \\ V(X) = & \; np(1 - p) \end{aligned}$$

Example. Flipping 3 times a coin.

X="number of heads" follows a $Bin\left(3, 0.5\right)$

$$E\left(X\right) = n \cdot p = 3 \cdot 0.5 = 1.5$$

- Probability of 2 heads

$$P\left(X = 2\right) = \binom{3}{2} 0.5^2 0.5 = 0.375$$

- Probability of less than 2 heads

$$P\left(X < 2\right) = P\left(X \leq 1\right) = F\left(1\right) = P\left(X = 0\right) + P\left(X = 1\right) = \\ \tbinom{3}{0} 0.5^0 0.5^3 + \tbinom{3}{1} 0.5^1 0.5^2 = 0.125 + 0.375 = 0.5$$

Geometric

X: "number of failures before a success".

Every trial follows an independent bernoulli distribution with parameter p.

- Range of values: $\{0, 1, 2, \ldots\}$
- Probability mass function

$$P\left(X = k\right) = \left(1 - p\right)^k p$$

- Probability distribution function

$$F\left(k\right) = 1 - \left(1 - p\right)^{k+1}$$

- Expectation and Variance

$$E\left(X\right) = \tfrac{1-p}{p} \qquad V\left(X\right) = \tfrac{1-p}{p^2}$$

Alternatively, one could define W="Number of trials until one success is observed"

$$W = X + 1$$

- Range of values: $\{1, 2, \ldots\}$
- Probability mass function

$$P\left(W = k\right) = \left(1 - p\right)^{w-1} p$$

- Probability distribution function

$$F(k) = 1 - (1-p)^k$$

- Expectation and Variance

$$E(X) = \tfrac{1}{p} \qquad V(X) = \tfrac{1-p}{p^2}$$

Example. W="Number of rolls of a die until one 5 is observed"

$$p = \frac{1}{6}$$

- Probability of (exactly) 6 rolls

$$P(W=6) = \left(\frac{5}{6}\right)^5 \frac{1}{6} = 0.067$$

- Probability of "number 5 is observed in 3 rolls as much"

$$P(W \leq 3) = F(3) = 1 - \left(1 - \frac{1}{6}\right)^3 = 0.421$$

- Expected number of rolls

$$E(W) = \left(\frac{1}{6}\right)^{-1} = 6$$

Negative Binomial

Number of failures until r successes are observed.

Generalization of the geometric distribution: sum of r independent and identical Geometric distributions.

- Range of values: $\{0, 1, 2, \ldots\}$
- Probability mass function

$$P(X=k) = \binom{k+r-1}{k} (1-p)^k p^r$$

$$E(X) = \frac{r(1-p)}{p}$$

$$V(X) = \frac{r(1-p)}{p^2}$$

Example. X = "Number of non-5 results until three 5 are observed when rolling a die"

$r = 3$, $p = 1/6$

$$X \sim NB\,(3, 1/6)$$

- Probability of 10 rolls (7 failures, 3 successes).

$$P\,(X = 7) = \binom{9}{7}\left(\frac{5}{6}\right)^7\left(\frac{1}{6}\right)^3 = 0.017$$

- Expected number of "non-5" results until three "5" are obtained.

$$E\,(X) = \frac{3\left(\frac{5}{6}\right)}{\frac{1}{6}} = 15$$

Poisson

Let X follow a Binomial distribution $Bin\,(n, p)$

$$E\,(X) = np$$

$$P\,(X = k) = \binom{n}{k} p^k\,(1-p)^{n-k} = \binom{n}{k}\left(\frac{\lambda}{n}\right)^k\left(1 - \frac{\lambda}{n}\right)^{n-k}$$

If the number of trials (n) tends to infinite and the probability of success (p) tends to zero in such a way that the expectation (λ) remains the same:

$$\lim_{n\to\infty} P\,(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}$$

In this case X follows a Poisson distribution of parameter λ

- Commonly applied to rare events.
- Usually X is expressed with relation to a differential (time, space,..).
- Range of values: $\{0, 1, 2, \ldots, \infty\}$
- Expectation and Variance

$$E\,(X) = V\,(X) = \lambda$$

Example

An emergency service from a certain hospital usually receives 10 visits per hour on average. The service collapses if it receives more than 20 visits in an hour.

- Probability that the service collapses.

$$P\left(X > 20\right) = 1 - P\left(X \leq 20\right) = 1 - \sum_{i=1}^{20} e^{-10} \frac{10^i}{i!} = 0.0016$$

- Probability of no visits in an hour.

$$P\left(X = 0\right) = e^{-10} \frac{10^0}{0!} = 4.54 \cdot 10^{-5}$$