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 males in sets of 3 siblings.

Assuming the same probability ofr male and female, $P\left(♀\right)=P\left(♂\right)=0.5$, and independence among births:

X = number of males.

The probability of each possible event (sample space), and the probability mass function of the variable "number of males" are:

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

Table 1: Probability of event

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

Table 2: Probability mass function

Probability Distribution Function

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

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

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

- Probability of an interval. Notice the sense of the inequality symbols!

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

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.

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

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

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

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

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

The expectation is 1.5 males. Note that the expected value may be a value that the random variable cannot take on.

Properties of the Expectation

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

- Linearity.

$$E\left(X + C\right) = E\left(X\right) + C$$
$$E\left(X + Y\right) = E\left(X\right) + E\left(Y\right)$$

- Multiplicity.

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

- If X and Y are independent: $E\left(X \cdot Y\right) = E\left(X\right) \cdot E\left(Y\right)$

Variance of a random variable

- Variance: it is a statistical distance that indicates the dispersion of the variable around its expectation.

The variance answers the following questions:

- 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 widely dispersed and the expectation will not accurately represent the random variable.

- General expression

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

- Variance in discrete variables

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

- Example: number of males in 3 siblings

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

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

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

Properties of the Variance

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

- Invariant to changes in location parameter

$$V(X + C) = V(X)$$

- Scale

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

- 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:

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

- Variance of the difference of two random variables

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

Probability distribution models

Bernoulli

- Binary results. The variable usually takes values 0 or 1.
- Probability distribution defined by one single parameter p (probability of success)
- Probability mass function

$$P(X=1)=p \quad P(X=0)=1-p$$

- 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 $Bin(3,0.5)$

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

- Probability of 2 heads

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

- Probability of less than 2 heads

$$P\left(X<2\right)=P\left(X\leq1\right)=F\left(1\right)=P\left(X=0\right)+P\left(X=1\right)=$$
$$\binom{3}{0}0.5^{0}0.5^{3}+\binom{3}{1}0.5^{1}0.5^{2}=0.125+0.375=0.5$$

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,

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
$$F\left(k\right)=1-\left(1-p\right)^{k}$$
- Expectation and Variance
$$E\left(X\right)=\tfrac{1}{p}\qquad V\left(X\right)=\tfrac{1-p}{p^{2}}$$

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

- Probability of 5 in one roll:

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

- Probability of (exactly) 6 rolls until one 5 is observed

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

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

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

- Expected number of rolls

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

Negative Binomial

X=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

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

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

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

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

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

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

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

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

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

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

- Expected number of rolls until three "5" are obtained.

The expected number of "non-5" results is 15, plus the three "5", the expected number of rolls is $15+3=18$.

Poisson

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

$$E\left(X\right)=np$$

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

Let's suppose the number of trials (n) tends to infinite and the probability of success (p) tends to zero in such a way that the expectation remains stable $np=\lambda$.

In this case, X follows a Poisson distribution of parameter λ with probability mass function:

$$P\left(X=k\right)=e^{-\lambda}\frac{\lambda^{k}}{k!}$$

Some properties:

- 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 $E\left(X\right)=V\left(X\right)=\lambda$

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.

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