1 Discrete Probability Distribution

1.1 Probability distribution

Random variable: A random variable, usually written x, is a variable whose possible values are numerical outcomes of a random phenomenon. There are two types of random variables, discrete and continuous.

The **probability distribution** $P(x)$ of a discrete **random variable** x is the probability that a certain value of x is realized. Note that $P(x)$ must have these properties:

$$0 \leq P(x) \leq 1 \quad \text{for all values of } x$$

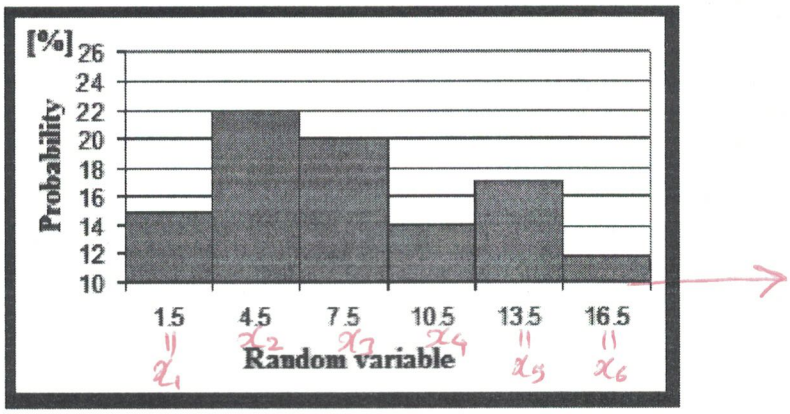
and

$$\sum_{i=1}^{N} P(x_i) = 1$$

where $\{x_1, x_2, x_3, \cdots, x_N\}$ represents *events* that are *disjoint* and their *union* is the *sample space*.

Example 1: The probability distribution of the number of calls per hour at a support line.

Discrete probability distribution: Histogram with probability on the vertical axis, In this case, the class mark of the histogram classes can be used as the random variable.



$$\sum_{i=1}^{6} P(x_i) = P(1.5) + P(4.5) + P(7.5) + P(10.5) + P(13.5) + P(16.5) =$$

$$= 0.15 + 0.22 + 0.20 + 0.14 + 0.17 + 0.12 = 1$$

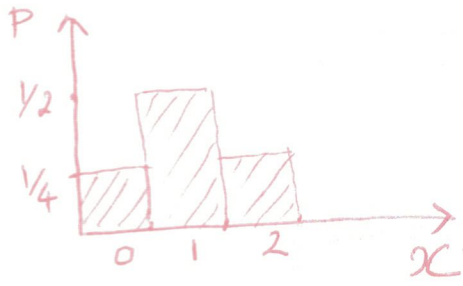
Example 2: 2 coin tosses. Find the probability distribution.
Random variable: $x =$ the number of heads: $\{0, 1, 2\}$

$S = \{(H,H), (H,T), (T,H), (T,T)\}$

$P(0) = \frac{1}{4}$

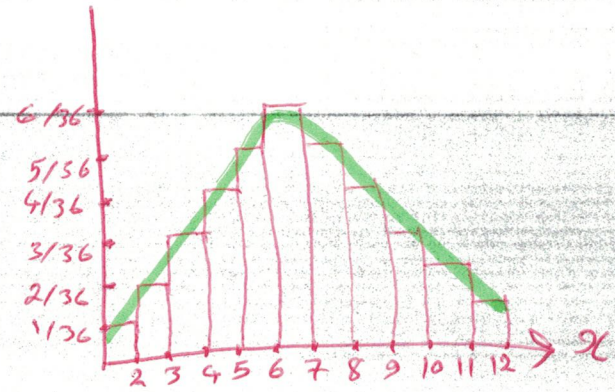
$P(1) = \frac{1}{2}$

$P(2) = \frac{1}{4}$



Example 3: Rolling 2 dice. Find the probability distribution.
Random variable: x = sum of the faces: $\{2, ..., 12\}$

$P(2) = P(12) = \frac{1}{36}$
$P(3) = P(11) = 2/36$
$P(4) = P(10) = 3/36$
$P(5) = P(9) = 4/36$
$P(6) = P(8) = 5/36$
$P(7) = 6/36$



1.2 Mean

Let us consider a very simple *toy model* $x = \{1, 1, 1, 2, 2, 2, 2, 5, 5\}$ (let's say this is the population). Then the mean is

$N = 9$

$$\mu = \frac{1}{9}(1+1+1+2+2+2+2+5+5) = \frac{1}{9}(\overbrace{3}^{f_1} \times 1 + \overbrace{4}^{f_2} \times 2 + \overbrace{2}^{f_3} \times 5) = \frac{1}{N}(f_1 x_1 + f_2 x_2 + f_3 x_3)$$

Thus, in general, if f_i is the frequency of each outcome x_i, the mean μ is

$$\mu = \frac{1}{N}\sum_i f_i x_i$$
$$= \sum_i \frac{f_i}{N} x_i = \sum_i P(x_i) x_i$$

$P(x_i) = \frac{f_i}{N}$

The mean is also called the **expected value** of x, and often denoted as

$$\mu = E(x) = \sum_i x_i P(x_i)$$

Example 4: The expected value of the number of heads in 2-coin tosses.

$$E(x) = \sum x_i P(x_i) = 0\,P(0) + 1 \cdot P(1) + 2 \cdot P(2)$$
$$= 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} = 1$$

The expected value of $g(x)$, a function of x, is

$$E(g(x)) = \sum_i g(x_i) P(x_i)$$

Example 5: The expected value of x^2 in the 2-coin toss experiment.

$$E(x^2) = \sum x_i^2 P(x_i) = 0^2 P(0) + 1^2 \cdot P(1) + 2^2 P(2)$$
$$= 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 4 \cdot \frac{1}{4} = 1.5$$

1.3 Standard Deviation

Again, using the frequencies f_i, the standard deviation can be expressed as

$$\sigma = \sqrt{\frac{1}{N}\sum_i f_i(x_i-\mu)^2} = \sqrt{\frac{1}{N}\Big[\underbrace{(x_1-\mu)^2+(x_1-\mu)^2+\cdots+(x_1-\mu)^2}_{f_1}+(x_2-\mu)^2+\cdots}$$

$$= \sqrt{\sum_i \underbrace{\frac{f_i}{N}}_{P(x_i)}(x_i-\mu)^2} = \sqrt{\sum_i P(x_i)\cdot\underbrace{(x_i-\mu)^2}} = \sqrt{E\left(\underbrace{\ ?\ }_{(x-\mu)^2}\right)}$$

$E(g(x)) = \sum g(x_i)P(x$

Example 6: Show that $\sigma = \sqrt{E(x^2)-\mu^2}$.

$$\sigma^2 = E((x-\mu)^2) = E(x^2-2\mu x+\mu^2)$$

$$= \sum (x_i^2-2\mu x_i+\mu^2)P(x_i)$$

$$= \sum x_i^2 P(x_i) - 2\mu\underbrace{\sum x_i P(x_i)}_{\mu} + \mu^2\underbrace{\sum P(x_i)}_{1}$$

$$= E(x^2) - 2\mu^2 + \mu^2 = E(x^2) - \mu^2$$

Example 7: Calculate the standard deviation for the 2-coin toss experiment.

$$\sigma = \sqrt{E(x^2)-\mu^2} = \sqrt{1.5-1^2} = \sqrt{0.5} \approx 0.7$$

2 Binomial distribution

Binomial distribution appears when

- the experiment consists of n identical trials
  (e.g., flip a coin n times, free throws in basketball game)
- the outcome is binary ("success" or "failure")
- the trials are independent
- the random variable x is the number of successes in n trials.

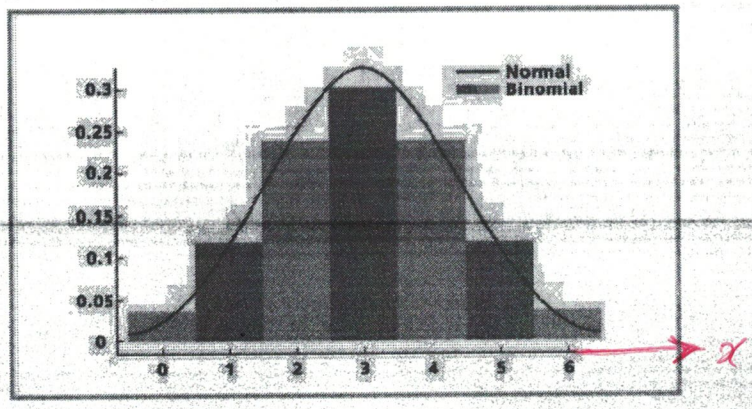
e.g., in 2 coin flip $n=2$

$x = \#$ of heads

6 Coin flip

$n = 6$

x = # of Heads



2.1 Probability function

The probability of having x successes in n trials is $P(x) = {}_nC_x\, p^x (1-p)^{n-x}$, where ${}_nC_x$ is the combination formula and p is the probability of success in **each** trial.

Example 8: If $n = 3$, calculate $P(x)$. $[0! = 1]$

$P(0) = C_{30}\, p^0 (1-p)^3 = \frac{3!}{(3-0)!\,0!}(1-p)^3 = (1-p)^3$ FFF

$P(1) = C_{31}\, p^1 (1-p)^2 = \frac{3!}{(3-1)!\,1!}\, p(1-p)^2 = 3p(1-p)^2$ SFF, FSF, FF

$P(2) = C_{32}\, p^2 (1-p) = \frac{3!}{(3-2)!\,2!}\, p^2(1-p) = 3p^2(1-p)$ SSF, SFS, FSS

$P(3) = C_{33}\, p^3 (1-p)^0 = \frac{3!}{(3-3)!\,0!}\, p^3(1-p)^0 = p^3$ SSS

2.2 Mean

The probability of "success" for each trial is p, thus, with n trials the expected value for the number of successes x is

$$\mu = E(x) = np$$

Example 9: The probability of a randomly chosen households having pet(s) is 45%. If the total number of households in Canada is 13 million. What is the expected number of households with pet(s) in Canada?

$n = 13{,}000{,}000$ $P = 0.45$

Having pets Success
No pets Failure
} $\longrightarrow$ Binomial Dist.

$\therefore \mu = np = 13{,}000{,}000 \times 0.45 = 5.85$ million

2.3 Standard deviation

The standard deviation of a binomial distribution is given as

$$\sigma = \sqrt{E((x-\mu)^2)} = \sqrt{np(1-p)}$$

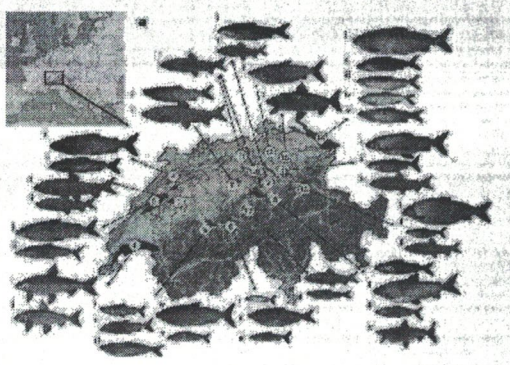
Example 10: Calculate the standard deviation for the experiment described in the previous example.

$n = 13 \times 10^6$ $p = 0.45$

$\sigma = \sqrt{np(1-p)} = \sqrt{(13\times10^6)(0.45)(1-0.45)}$

$= 10^3\sqrt{13\times(0.45)\times(0.55} = 1794$

3 Poisson Distribution



In Poisson distribution, the random variable is the number of occurrences of some event (e.g., earthquakes, telephone calls) over a specified interval.

3.1 Probability function

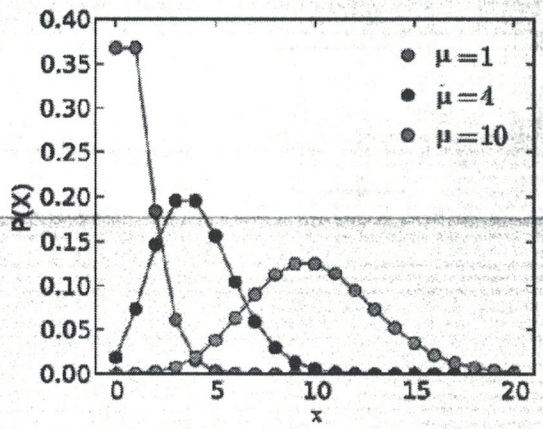
The Poisson distribution is the binomial distribution, except for the fact that it gives the probability of occurrence over a specified interval in time or space. Thus, the Poisson distribution can be derived from the binomial distribution:

1. The expected value (mean) over a fixed interval is given, e.g. on the average 96 large earthquakes occur every 100 years. ($\mu = 96$)
2. Divide the interval (100 years) into n small subintervals of equal length. If the interval is small enough, the event either occurs once or does not occur at all. Then, the distribution of the number of occurrences should be described by the binomial distribution.
3. Let us denote the probability of occurrence at each subinterval $p = \frac{\mu}{n}$, then

$$P(x) = {}_nC_x\, p^x (1-p)^{n-x} = \frac{n!}{(n-x)!x!}\left(\frac{\mu}{n}\right)^x \left(1-\frac{\mu}{n}\right)^{n-x}$$

Taking $n \to \infty$, hence very small p, to represent a continuous interval, we obtain the Poisson distribution

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



3.2 Standard deviation

The standard deviation of the Poisson distribution is

$$\sigma = E((x-\mu)^2) = \sqrt{\mu}$$

Example 11: Find the probability of having exactly 10 major earthquakes in the 21st century.

$\mu = 96$

$x = 10$

$$P(x) = \frac{\mu^x e^{-\mu}}{x!} = \frac{(96)^{10} e^{-96}}{10!}$$

$$= 3.72 \times 10^{-29}$$

Prob of having 2 or more earthquakes?

$$P(x \geq 2) = P(2) + P(3) + P(4) + \cdots$$

$$= 1 - P(0) - P(1) = 1 - P(x \leq 1)$$

$$= 1 - \frac{96^0 e^{-96}}{0!} - \frac{96^1 e^{-96}}{1!}$$