Probability and Statistics - Elementary Probability Theory

Giuliano Casale
Department of Computing, Imperial College London

Expectation of a Sum of Random Variables

Let $X_1, X_2, \ldots, X_n$ be n random variables, possibly with different distributions and not necessarily independent.

Let $S_n = \sum_{i=1}^{n} X_i$ be their sum, and $\bar{X} = \frac{S_n}{n}$ be their average.
Then:

$$\mathrm{E}(S_n) = \sum_{i=1}^{n} \mathrm{E}(X_i), \quad \mathrm{E}(\bar{X}) = \frac{\sum_{i=1}^{n} \mathrm{E}(X_i)}{n}.$$

We will give a proof when we consider joint random variables.

Variance of a Sum of Independent Random Variables

If the random variables $X_1, X_2, \ldots, X_n$ are independent, then:

$$\mathrm{Var}(S_n) = \sum_{i=1}^{n} \mathrm{Var}(X_i), \quad \mathrm{Var}(\bar{X}) = \frac{\sum_{i=1}^{n} \mathrm{Var}(X_i)}{n^2}$$

So if $X_1, X_2, \ldots, X_n$ are independent and identically distributed (i.i.d.) with $\mathrm{E}(X_i) = \mu_X$ and $\mathrm{Var}(X_i) = \sigma_X^2$ we get

$$E(\bar{X}) = \mu_X, \quad \mathrm{Var}(\bar{X}) = \frac{\sigma_X^2}{n}.$$

Notable Discrete Distributions

Bernoulli(p)

- Consider an experiment with only two possible outcomes, encoded as a random variable X taking values 1 , with probability p; and 0 , with probability ($1 - p$).

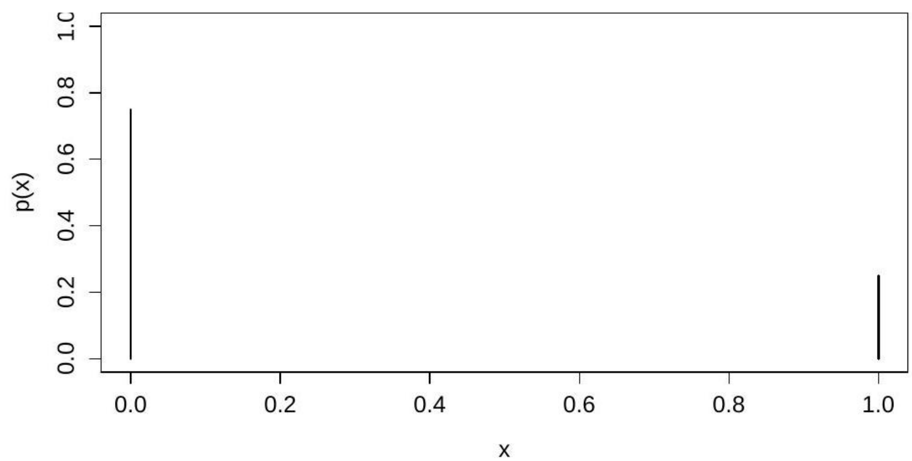
- For example, tossing a coin with probability p for heads: $X = 1$ for heads; $X = 0$ for tails.
- Then we say $X \sim \text{Bernoulli}(p)$ and note the pmf to be

$$p(x) = p^x (1-p)^{1-x}, \quad x = 0, 1.$$

- Using the formulae for mean and variance, it follows that

$$\mu = p, \quad \sigma^2 = p(1-p).$$

Example: Bernoulli $\left(\frac{1}{4}\right)$ pmf



Binomial(n, p)

- Now consider n identical, independent Bernoulli (p) trials $X_1, \ldots, X_n$.
- Let $X = \sum_{i=1}^{n} X_i$ be the total number of 1 s observed in the n trials.
- For example, tossing a fair coin n times, X may be the number of heads obtained, $p = \frac{1}{2}$.
- Then X is a random variable taking values in $\{0, 1, 2, \ldots, n\}$, and we say $X \sim \text{Binomial}(n, p)$.
- From the Binomial Theorem we find the pmf to be

$$p(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, 2, \ldots, n.$$

Proof and moments

- Use simple combinatorial arguments, remembering that

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}$$

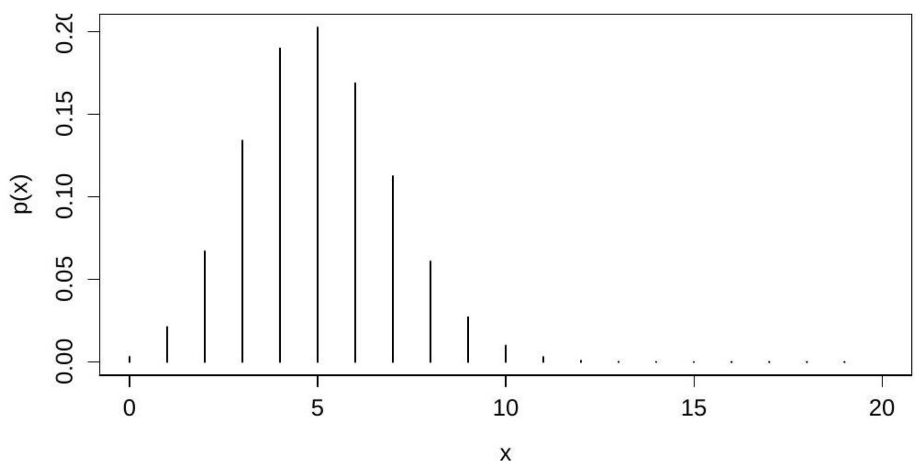
- Mean and variance are (from pmf or results on sums of r.vs.)

$$\mu = np, \quad \sigma^2 = np(1-p)$$

- Similarly, the skewness is:

$$\gamma_1 = \frac{1-2p}{\sqrt{np(1-p)}}$$

Example: Binomial $(20, \frac{1}{4})$ pmf



Geometric(p)

Consider a potentially infinite sequence of independent Bernoulli (p) random variables $X_1, X_2, \ldots$.

- Suppose we define a quantity X by

$$X = \min\{i \mid i \geq 1, X_i = 1\}$$

to be the index of the first Bernoulli trial to result in a 1 .

- Then X is a random variable taking values in $\text{supp}(X) = \{1, 2, \ldots\}$, and we say $X \sim \text{Geometric}(p)$.

Example: Tossing a coin

- X is the number of tosses until the first head is obtained.
- The pmf is:

$$p(x) = p(1-p)^{x-1}, \quad x = 1, 2, \ldots$$

- The mean and variance are:

$$\mu = \frac{1}{p}, \quad \sigma^2 = \frac{1-p}{p^2}$$

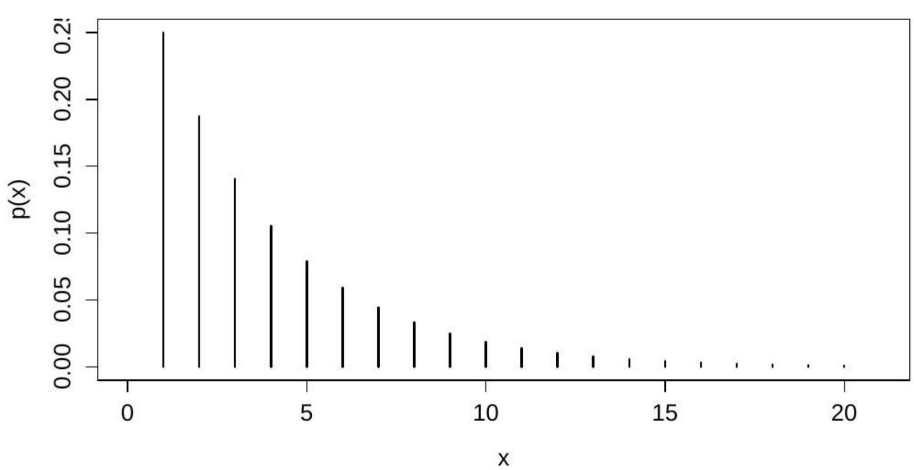
- The skewness is:

$$\gamma_1 = \frac{2-p}{\sqrt{1-p}}$$

and so is always positive.

Remark: some texts call Geometric the distribution of the number of trials before we obtain our first 1 . Formulas for pmf and moments are similar.

Example: Geometric $\left(\frac{1}{4}\right)$ pmf



Poisson: Poi(λ)

- Let X be a random variable on $\mathbb{N} = \{0, 1, 2, \ldots\}$. Define

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

for some $\lambda > 0$.

- Then, X is said to follow a Poisson distribution with parameter λ and we write $X \sim \text{Poi}(\lambda)$.
- Poisson random variables are concerned with the number of random events occurring per unit of time or space, if there is a constant rate of random events occurring across this unit.

Examples

- the number of patients arriving at an emergency room in a hour;
- the number of minor car crashes per day in the U.K.;
- the number of potholes in each mile of road;
- the number of jobs which arrive at a database server per hour;
- the number of particles emitted by a radioactive substance in a given time.
- An interesting property of the Poisson distribution is that it has equal mean and variance, namely

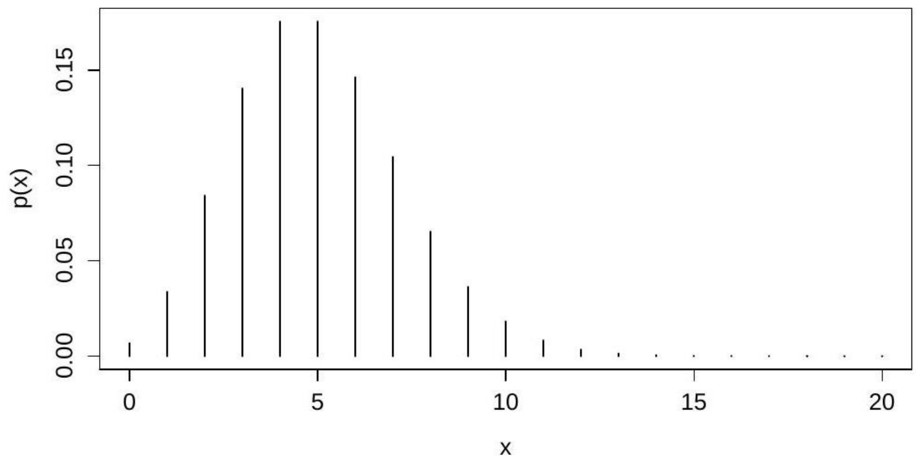
$$\mu = \lambda, \quad \sigma^2 = \lambda$$

- The skewness is given by

$$\gamma_1 = \frac{1}{\sqrt{\lambda}}$$

so is always positive but decreasing as λ grows.

Example: Poi(5) pmf



Using the Poisson Distribution in practice

- What do we do if we have a non-unit interval (or space) of length t ?
- In this case, λt can be used in the pmf instead of λ, so that

$$p(x) = \frac{e^{-\lambda t}(\lambda t)^x}{x!}, \quad x = 0, 1, 2, \ldots$$

and we write $X \sim \text{Poi}(\lambda t)$.

- We thus now see λ as the rate at which random events occur and λt as the mean number of events in t.
- Thus, both with $t = 1$ and $t \neq 1$, we see the input parameter to the Poisson as the mean of the distribution.

Uniform: $\cup(\{1, 2, \ldots, n\})$

- Let X be a random variable on $\{1, 2, \ldots, n\}$ with pmf

$$p(x) = \frac{1}{n}, \quad x = 1, 2, \ldots, n.$$

- Then X is said to follow a discrete uniform distribution and we write $X \sim \text{U}(\{1, 2, \ldots, n\})$.
- The mean and variance are

$$\mu = \frac{n+1}{2}, \quad \sigma^2 = \frac{n^2-1}{12}.$$

- Q: what value do you expect for the skewness?

Q&A: Coupon collector problem

A company producing cereals places 1 coupon in each cereal box. There are m types of coupons and we wish to collect them all.

Suppose that each coupon type is equally-likely and independent of what has been previously obtained, i.e., the cereal boxes on the market are so many that drawing can be assumed with replacement.

Q: Find the mean number of boxes X that we need to obtain in order to have at least one coupon of each type.