No more than 2 successes in first 30 trials. [See the 3rd success from 31 on]

① Define number of trials as R.V.

then $Y \sim nb(3, 0.15)$

$P(Y \geq 31) = 1 - P(30) = 1 0.151$

② If we define successes as R.V. and think $n = 30$ is fixed number of trials

then $Y \sim b(30, 0.15)$

$P(Y=1) + P(Y=0) + P(Y=2)$

$= 0.151.$

Two different ways to approach this problem.

Go back to hyper-geometric notes.

Recap. ① Four candidates.

A B C. D

Choose 2 from them [same positions]

How to do that

AB, AC, AD, BC, BD, CD

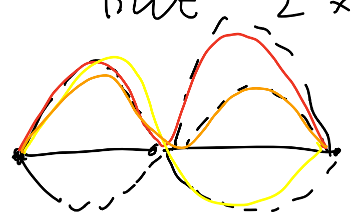
or $\binom{4}{2} = \frac{4!}{2!\,2!}$

$= \frac{4\times3\times2\times1}{2\times1\times2\times1} = \frac{4\times3}{2} = 6$

② If it takes two steps to finish one job, and there are 2 & 3 different ways to finish for each step, respectively. Then, how many different ways can ~~way~~ we chose ~~from entire~~ for the entire job?

NOT 2 + 3

But, 2 × 3



Example 3.7. a supplier ships parts to a company in lots of 100 parts. The company has an acceptance sampling plan which adopts the following acceptance rule:

"sample 5 parts at random and with no replacement"

If there are no defective in the sample, accept the entire lot, Otherwise, reject the entire lot.

The population size is $N = 100$, the sample size is $n = 5$. Define the random variable.

Y = the number of defectives in the sample $\sim hyper(N=100, n=5, r)$

(a) If there are 10 defectives in the population, what is the probability that the lot will be accepted.

$$P_r(0) = P(Y=0) = \frac{\binom{10}{0}\binom{90}{5}}{\binom{100}{5}} \approx 0.584$$

(b) If $r=10$, what is the prob at least 3 of the 5 parts sampled are defective

$$P(Y \geq 3) = 1 - P(Y \leq 2)$$

$$= 1 - [P(Y=0) + P(Y=1) + P(Y=2)]$$

$$= 1 - \left[\frac{\binom{10}{0}\binom{90}{5}}{\binom{100}{5}} + \frac{\binom{10}{1}\binom{90}{4}}{\binom{100}{5}} + \frac{\binom{10}{2}\binom{90}{3}}{\binom{100}{5}} \right]$$

$$= 1 - (0.584 + 0.339 + 0.070)$$

$$\doteq 0.007$$

R code for hyper-geometric

$p_Y(y) = P(Y=y) =$ dhyper(y, r, N-r, n)

$F_Y(y) = P(Y \leq y) =$ phyper(y, r, N-r, n)

part (1) dhyper(0, 10, 100-10, 5) = 0.584

part (2) 1-phyper(2, 10, 100-10, 5) = 0.0066

3.6 poisson distribution

Note: The poisson distribution is commonly used to model counts. such as

1. the # of customers entering post office in a given hour
2. The number of machine Breaks down per month
3. # of insurance claims received per day

In general, we define

$Y =$ # of "occurrences" over a unit interval of time (or space)

A poisson distribution for Y emerges if "occurrences" obey the following assumptions:

① The number of occurrences in non overlapping intervals are independent.

② The ~~pro~~ probability of an occurrence is proportional to the length of the interval

③ The probability of 2 or more occurrences in a sufficiently short interval is 0.

we say that Y has a Poisson distribution

$$Y \sim \text{Poisson}(\lambda)$$

A processes that produces occurrences according to these assumptions is call poisson proccess.

pmf If $Y \sim \text{Poisson}(\lambda)$, then the pmf of Y is given by

$$p_Y(y) = \begin{cases} \frac{\lambda^y e^{-\lambda}}{y!} & y = 0, 1, \ldots \\ 0, & \text{o.w} \end{cases}$$

Mean / variance

If $Y \sim \text{Poisson}(\lambda)$, then ~~E(λ)~~

$$E(Y) = \lambda \; ; \; Var(Y) = \lambda$$

Example 3.8.

Let Y denote the number of time per month that a detectable amount of radioactive gas is recorded at a nuclear power plant. Suppose that Y ~~has~~ follows a Poisson distribution with $\lambda = 2.5$ times per month.
mean

(a). What is the probability that there are exactly 3 times that a detectable amount of gas is recorded in a given month.

$$P(Y=3) = p_Y(3) = \frac{(2.5)^3 e^{-2.5}}{3!}$$

$$= \frac{15.625 \times e^{-2.5}}{6}$$

$$= 0.214$$

(b) What is the probability that there are no more than 4 times a detected amount of gas is recorded in a month?

$$P(Y \le 4) = P(Y=0) + P(Y=1) + P(Y=2) + P(Y=3) + P(Y=4)$$

$$= \frac{(2.5)^0 e^{-2.5}}{0!} + \frac{(2.5)^1 e^{-2.5}}{1!}$$

$$+ \frac{(2.5)^2 e^{-2.5}}{2!} + \frac{(2.5)^3 e^{-2.5}}{3!}$$

$$+ \frac{(2.5)^4 e^{-2.5}}{4!} \approx 0.891$$

~~Pos~~ Poisson R code

$$p_Y(y) = P(Y = y) = \text{dpois}(y, \lambda)$$

$$F_Y(y) = P(Y \le y) = \text{ppois}(y, \lambda)$$

part (a) = dpois (3, 2.5) = 0.213763

part (b) = ppois (~~3~~ 4, 2.5) = 0.891178

Chapter 4,

Continuous distribution

Recall:

A r.v. Y is called continuous if it can assume any value in an interval of real numbers.

- Contrast this with a random discrete variable who values can be counted.

Important

Assigning probabilities to event involving continuous R.V is different than in discrete models. We don't assign positive probability to specific values (e.g. $Y = 3$, ~~etc~~ etc) like we did in discrete R.V.s.

Instead, we assign positive ~~probabilities~~ probability to events which are intervals (e.g. $2 < Y < 4$ etc.)

pdf

Every continuous r.v. we will discuss in this course has a pdf (probability ~~distribution~~ density function), denoted by $f_Y(y)$. The function has the following ~~characters~~ characteristics:

1. $f_Y(y) \geq 0$, that is $f_Y(y)$ is non-negative

2. The area under any pdf is equal to 1, that is

$$\int_{-\infty}^{+\infty} f_Y(y)\,dy = 1$$

CDF

~~$P(Y \leq$ $F_Y(y) =$~~

The cumulative distribution function of Y is given by

$$F_Y(y) = P(Y \le y) = \int_{-\infty}^{y} f_Y(t)\,dt$$

Result.

If a and b are specific values of interest ($a \le b$), then

$$P(a \le Y \le b) = \int_a^b f_Y(y)\,dy$$

$$= F_Y(b) - F_Y(a)$$

Result.

If a is a specific values then $P(Y = a) = 0$

In other words, in continuous probability models, specific points are assigned 0 probability. An immediate consequence of this is that if Y is continuous,

$$P(a \le Y \le b) = P(a \le Y < b)$$

$$= P(a < Y \le b) = P(a < Y < b)$$

And each equal to

$$\int_a^b f_Y(y)\,dy$$

This is not true if Y has a discrete distribution. since positive probability is assigned to specific values of Y.