Discrete random variables

DRV

Continous data

can take any Value

$x = 1, 1.3, 2.4$

ex $[1, 10]$

$1 \leq x \leq 10$

"measured"

temperature

$\{83.6^{\circ}F, 99.46^{\circ}\}$

Discrete data

- only have specific Value
- finite

$x = 1, 2, 3$

"counted"

die

$\{1, 2, 3, 4 \ldots\}$

Random Variable

RV Variable whose value is determined by a random experiment

Discrete prob Dist

Table or formula that list the probabilities for each outcome of the random Variable X → Capital X means random. Variable

flip 3 Coins at same time

let rv X be the # of heads Showing

$\boxed{x \rightarrow \text{means outcome}}$

Discrete prob Distribution

H H H → 3 heads
H H T → 2 heads
H T H → 2 heads
H T T → 1 head
T H H → 2 head
T H T → 1 head
T T H → 1 head
T T T → 0 head

x	0	1	2	3	H
$P(X=x)$	1/8	3/8	3/8	1/8	

prob that rv (X) = outcome (x)

ex prob dist for X, sum of 2 rolled dice

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

$2 \to 1+1$ $\quad 3 \to \begin{matrix} 1+2 \\ 2+1 \end{matrix}$

$\boxed{\text{Sum of all probabilitie} = 1 \text{ (always)}}$

Some of all Probapility = 1

$$\sum_{i=0}^{3} P(X=i) = 1 \longrightarrow P(\Omega) = 1$$

$$P(X \le 2) = P(X=0 \text{ or } X=1 \text{ or } X=2)$$

$$= P(X=0) + P(X=1) + P(X=2) = 7/8$$

or

$$P(X \le 2) = 1 - P(X=3) = 1 - \frac{1}{8} = 7/8$$

Probability Mass Function $\boxed{PMF}$

$f(x)$ $\quad X$ = # heads if you flip coin 2

$x = \{0, 1, 2\}$

PMF ← what the prob of getting 1 head

* $f(1) = P(x=1) = P(h,t) + P(t,h)$
$= \frac{1}{4} + \frac{1}{4} = \frac{2}{4}$ or 0.5

* $f(0) = P(x=0) = P(t,t) = \frac{1}{4}$

* $f(2) = P(x=2) = P(h,h) = \frac{1}{4}$

	1	2	3
x	0	1	2
$f\,x$	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$

$\rightarrow$ $n = 3$

$\boxed{\text{Rules}}$

1) $f(x_i) \geq 0$ $\quad$ must be greater than 0
$\quad x_i \longrightarrow 0, 1, 2$ or any # in se

2) $\sum_{i=1}^{n} f(x) = 1$

For every possible value x of the random variable (X). The Pmf specifies the Prob. of observing that value when the experiment is performed.

in some cases P overlaps making more than 1 (but dont worry about it too much)

Ex:

DRV $x \in \{0, \pm 1, \pm 2, \pm 3, \dots\}$

Given: $P(X \le 15) = 0.3$, $P(15 < X \le 24) = 0.6$, $P(X > 20) = 0.5$

a) $P(X > 15) = P(X \in \{16, 17, \dots\}) = 1 - P(X \le 15) = 0.7$

b) $P(X \le 24) = P(X \le 15) + P(15 < X \le 24) = 0.9$

c) $P(15 < X \le 20) = P(X > 15) - P(X > 20) = 0.2$

Cumulative Distribution function $\boxed{\text{cdf}}$

Cumulative Distribution Functions

Cumulative distribution functions (cdf) are used to calculate the area under the curve to the left from a point of interest. It is used to evaluate the accumulated probability. For continuous probability distributions, the probability = area under the curve. Total Area = 1.

The probability density function (pdf) is f(x) which describes the shape of the distribution. (uniform, exponential, or normal distribution)

$$\boxed{\text{cdf} = P(X \le x) = \frac{x - a}{b - a}}$$

$$\text{pdf} = f(x) = \frac{1}{b-a}$$

Ex: $x \in \{0, 1, 2, 3\}$

PMF

Pmf

x	0	1	2	3
$P(X=x)$	0.3	0.2	0.1	0.4

cdf

$x \le 1 = P(x \le 1)$

0+1

cdf

x	0	1	2	3
$F(x)$	0.3	0.5	0.6	1

$P(X \le 1) = P(X=0) + P(1) = 0.3 + 0.2 = 0.5$

Expected value of DRV

The expected value of a random variable is the theoretical mean of the random variable

To calculate the expected value for a discrete random variable X:

$$\boxed{E(X) = \sum_{\text{all } x} x P(x)}$$

$$\boxed{E(X) = \mu} = \text{mean}$$

Expectation of X

all values of x multiplied by their prob and summed

Expectation of a function g(X): $E[g(X)] = \sum_{\text{all } x} g(x)\, p(x)$

Var(X) is σ^2

Variance of X: $E[(X-\mu)^2] = \sum_{\text{all } x} (x-\mu)^2 p(x)$

Var(x)

$$\overbrace{E[(X-\mu)^2]}^{\sigma^2} = \overbrace{E(X^2)}^{E(X^2)} - \overbrace{[E(X)]^2}^{\mu^2}$$

Suppose 60% of American adults approve of the way the president is handling his job.

Randomly sample 2 American adults.

Let X represent the number that approve.

$X = 0, 1, 2$

60% (support) = 0.6

40% (NOS) = 0.4

Value of X:	SS	SN	NS	NN
	2	1	1	0
Prob	0.6 × 0.6	0.6 × 0.4	0.4 × 0.6	0.4 × 0.4

The Prob dist of X:

x	0	1	2
$P(x)$	0.16	0.48	0.36

or

$E(x^2) - (E(X))^2$

$1.92 - 1.2^2 = 0.48$

$$\frac{E(X)}{\mu} = \sum_{\text{all } x} x\,p(x) = 0 \times 0.16 + 1 \times 0.48 + 2 \times 0.36$$

$$= \boxed{1.2}$$

$$E(X^2) = 0^2 \times 0.16 + 1^2 \times 0.48 + 2 \times 0.36 = \boxed{1.92}$$

$\sigma^2 = 0.48$ standard divation $\sigma = \sqrt{\sigma^2}$

$$Var(x) = E[X-\mu)^2] = \sum_{\text{all } x} (x-\mu)^2 p(x)$$

$$= (0-1.2)^2 \times 0.16 + (1-1.2)^2 \times 0.48 + (2-1.2)^2 \times 0.36 = 0.48$$

Ex: 3 laptops

X = # laptops sold (DRV)

cost = \$500/each

Price = \$1000/each

Salvage Value = \$200/each

x	0	1	2	3
$P(X=x)$	0.1	0.2	0.3	0.4

$E[X] = 2$

$0\times0.1 + 1\times0.2 + 2\times0.3 \times 3\times0.4 = 2$

Y = total profit; $Y(x) = \underset{\text{selling price}}{1000} \cdot \underset{\text{\# sold}}{x} - \underset{\text{cost of lap}}{500} \cdot \underset{\text{\# of laptop}}{3} + \underset{\text{Salvag Value}}{200}\underbrace{(3-x)}_{\text{\# of not lap sold}}$

$Y(x) = 800x - 900$

y	-900	-100	700	1500
$P(Y=y)$	0.1	0.2	0.3	0.4

DRV of Y

$E[Y] = -900\cdot0.1 + (-100)\cdot0.2 + 700\cdot0.3 + 1500\cdot0.4 = \boxed{700}$

1. $E[x^2] = 0^2\cdot0.1 + 1^2\cdot0.2 + 2^2\cdot0.3 + 3^2\cdot0.4 = 5$

2. $Var(x) = E[x^2] - E^2[x] = 5 - 2^2 = 1$

3. $\sigma_x = \sqrt{1} = 1$

Bernoulli Distribution

- Suppose we have a single trial
- The trial can result in one of two possible outcomes
- P(success) = p
- P(Failure) = 1-p
- Let $X=1$ if success occurs
- Let $X=0$ if failure occurs

- If all these conditions are met then X has a Bernoulli Dist

$$\boxed{P(\overset{P(X)}{X} = x) = p^x(1-p)^{1-x}}$$

for $X = 0, 1$

$$P(X=1) = p^1 (1-p)^{1-1} = p$$

$$P(X=0) = p^0 (0-p)^{1-0} = 1-p$$

mean

$$\boxed{\mu = p}$$

Var(X)

$$\boxed{\sigma^2 = p(1-p)}$$

Approximately 1 in 200 American adults are lawyers.

One American adult is randomly selected.

What is the distribution of the number of lawyers?

Bernoulli with $p = \frac{1}{200}$

$$P(X = x) = \left(\frac{1}{200}\right)^x \left(1 - \frac{1}{200}\right)^{1-x}$$

for $x = 0, 1$

$$P(X=1) = \frac{1}{200} \qquad P(X=0) = \frac{199}{200}$$

Geometric RV

* Geometric dist is the dist of the number of trials needed to get the first success in repeated Bernoulli trials

- meaning there is independent trials, that result in either success or failure

- X represents the # trials needed to get first success

pmf formula

- For the first timer success to occur on the xth trial:

1- the first X-1 trials must be a failures $(1-p)^{x-1}$

2- the xth trial must be a success p

$$\boxed{P(X=x) = (1-p)^{x-1} p}$$

$\mu = \frac{1}{p}$ | $\sigma^2 = \frac{1-p}{p^2}$ | $P(X=x) = (1-p)^{x-1}p$

In a large population of adults, 30% have received CPR training.

If adults from this population are randomly selected, what is the probability that the 6th person sampled is the first that has received CPR training?

$p = 0.3$

$P(X=6) = (1-0.3)^{6-1}\,0.3$

$= \boxed{0.0504}$

$\mu = \frac{1}{0.3} = 3.3$ $\quad \sigma^2 = \frac{1-0.3}{0.3^2} = 7.7$

standard deviation $\sigma = 2.79$

* Prob the first person occurs on or before the 3rd person sampled?

$P(X \leq 3) = P(X=1) + P(X=2) + P(X=3)$

$= (1-0.3)^{1-1}0.3 + (1-0.3)^{2-1}0.3 + (1-0.3)^{3-1}0.3$

$= 0.7^0\,0.3 + 0.7^1\,0.3 + 0.7^2\,0.3$

$= \boxed{0.657}$

* comulative dist function

$\boxed{F(x) = P(X \leq x) = 1-(1-p)^x}$ cdf

ex) $P(X>3) = 0.7^3$

$P(X \leq 3) = 1 - 0.7^3 = 0.657$

Bionomial RV and dist

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

The # of successes in n independent Bernoulli trials has a binomial dist

There are n independent trials that has two outcomes either succeed or failure

$S = p \quad F = 1-p$

X represents the # of successes in n trials

pmf

$n \to$ # of trials

$$\boxed{P(X \overset{P(x)}{=} x) = \binom{n}{x} p^x (1-p)^{n-x}}$$

$x =$ # of success

$$\boxed{\mu = E(X) = np}$$

$p \to$ success

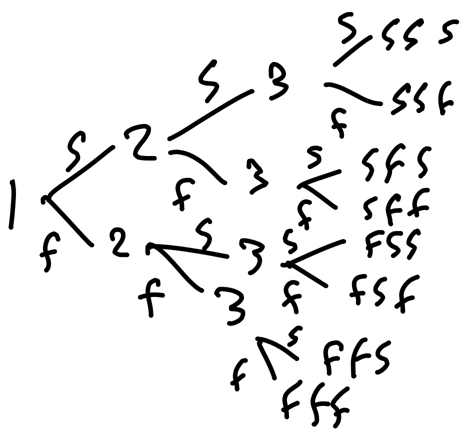
$$\boxed{\sigma^2 = np(1-p}$$

A balanced, six-sided die is rolled 3 times. What is the probability a 5 comes up exactly twice?

sucess = rolling a 5

$n = 3 \qquad p = 1/6$

$$P(x=2) = \binom{3}{2}\left(\frac{1}{6}\right)^2\left(1-\frac{1}{6}\right)^{3-2} = \boxed{0.0694}$$



According to Statistics Canada life tables, the probability a randomly selected 90 year-old Canadian male survives for at least another year is approximately 0.82.

If twenty 90-year old Canadian males are randomly selected, what is the probability exactly 18 survive for at least another year?

success $\to$ survive for a year

$n = 20 \qquad p = 0.82 \qquad X \sim B(20, 0.82)$

$$P(X=18) = \binom{20}{18} 0.82^{18} (1-0.82)^{20-18} = \boxed{0.173}$$

$$\mu = np = 16.4 \qquad \sigma^2 = np(1-p) = 2.952$$

$$\sigma = \sqrt{2.952}$$

at lest 18 survive

$$P(X \geq 18) = P(X=18) + P(X=19) + P(X=20)$$
$$= 0.173 + 0.083 + 0.019$$
$$= \boxed{0.275}$$

negative Binomial RV

* Geometric dist is the dist of the number of trials needed to get the first success in repeated Bernoulli trials

The negativ binomial dist is the dist of the number of trials needed to get the rth success

if wanted the 2nd sucess $r=2$
12 success $r=12$

Binomial

The # of successes X in a fixed number n of independent Bernoulli trials has a binomial dist

vs

negative Bi

The negative binomial dist is the dist of the # of trials needed X to get a fixed number of success r

or # of failures to get X

There are independent trials that has two outcomes either succeed or failure

$S = p$
$f = 1-p$

X represents the trials # of the rth success

The first x-1 trial results in r-1 success

$$\binom{x-1}{r-1} p^{r-1} (1-p)^{(x-1)(r-1)} \leftarrow \text{Binomial formula}$$

The xth trial must be a success. Which has a prop of p

$\binom{x-1}{r-1} p^r (1-p)^{x-r}$ | $\mu = \frac{r}{p}$ | $\sigma^2 = \frac{r(1-p)}{p^2}$

pmf ↑

Pmf: $P(X=x) = \binom{x+r-1}{r-1} p^r (1-p)^x$ in teacher note — Use this one

Example: A person conducting telephone surveys must get 3 more completed surveys before their job is finished.

On each randomly dialed number, there is a 9% chance of reaching an adult who will complete the survey.

What is the probability the 3rd completed survey occurs on the 10th call?

$P(X=10)$ $p = 0.09$

$r = 3$

$$P(X=10) = \binom{10-1}{3-1} 0.09^3 (1-0.09)^{10-3} = 0.013$$

a) X = # of boxes that do not contain a prize until you find 2 prizes

$p = 0.2$
$r = 2$

$$P(X=x) = \binom{x+r-1}{r-1} p^r (1-p)^x = \boxed{\binom{x+1}{1} 0.2^2 (0.8)^x}$$

b) P(Purchase 4 boxes) = P(2 boxes w/o prizes)

$$P(X=2) = \binom{3}{1} 0.2^2 \cdot (0.8)^2 = \boxed{0.0768}$$

$$\frac{3!}{(3-1)!\,1!}$$

c) P(Purchase at most 4 boxes)

$$P(X \le 2) = P(X=0) + P(X=1) + P(X=2)$$
$$= \sum_{x=0}^{2} \binom{x+1}{1} (0.2)^2 (0.8)^x = \boxed{0.1806}$$

d) $E[X] = \frac{r(1-p)}{p} = \frac{2(1-0.2)}{0.2} = \boxed{8}$

In total $\Rightarrow$ $E[X] + 2 = 10$

imp

Expected value + 2 prizes = 10 trials

Hyper Geometric dist

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

We are randomly sampling n objects without replacement from a source that contains a successes and N- a ailures

X represents the number of successes in the sample

X has the hypergeometric dist: $$\boxed{P(X = x) =} \frac{\binom{a}{x}\binom{N-a}{n-x}}{\binom{N}{n}}$$

- teacher uses M instead of a

$$\boxed{\mu = n\frac{a}{N}}$$ or $n\frac{M}{N}$ from notes

$$\boxed{\sigma^2 = \left(\frac{N-n}{n-1}\right) n p (1-p)}$$

Suppose a large high school has 1100 female students and 900 male students.

A random sample of 10 students is drawn. What is the probability exactly 7 of the selected students are female?

$$P(X=7) = \frac{\binom{1100}{7}\binom{900}{3}}{\binom{2000}{10}} = \boxed{0.166490}$$

$N = 2000$ $\quad a = 1100$

$n = 10$

Poisson Dist.

Events are occurring independently

The prob that an event occurs in a given length of time does not change the time (the theoretical rate does not change through time)

Then X the number of events in a fixed unit of time has a poison dist

$$\boxed{P(X=x)=\frac{\lambda^{x}e^{-\lambda}}{x!}}$$

pmf

$$\text{Pmf}: P(X=k)=\frac{e^{-\mu}\cdot\mu^{k}}{k!},\quad k=0,1,2,\ldots$$

$$E[X]=\mu \qquad Var(X)=\mu$$

$\lambda=\mu$

$\sigma^2=\lambda$

$e\approx 2.71828$

One nanogram of Plutonium-239 will have an average of 2.3 radioactive decays per second, and the number of decays will follow a Poisson distribution.

What is the probability that in a 2 second period there are exactly 3 radioactive decays?

$x=2$

$\lambda=2.3\times 2=4.6$

$$P(x=3)=\frac{\lambda^{x}e^{-\lambda}}{x!}=\frac{4.6^{3}\,e^{-4.6}}{3!}=\boxed{0.163}$$

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

if n is large and p is small in $\lambda=np$ it's most likely a possion

Continuous Random variable

Continuous random variables can take on an infinite number of possible value. (3, 3.1, 3.12...,4) or $(0,\infty)$

We cannot model continuous random variables with the same methods we used for discrete random variable

We model continuous random variables with a curve f(X), called probability density function (pdf)

- For continuous random variables, probabilities are areas under the curve

$$\boxed{P(a\leq X\leq b)=\int_a^b f(x)\,dx}$$

* $P(X=a)=0$
* $f(x)\geq 0$
* Area under entire curve equal to one

Continuous prob dist

$$\boxed{\int_{-\infty}^{\infty} f(x)\,dx = 1}$$

Suppose for a random variable X:

$$f(x) = cx^3$$

for $2 \le x \le 4$ and 0 otherwise.

What value of c makes this a legitimate probability distribution?

$$\int_2^4 cx^3\,dx = 1$$

$$\int_2^4 cx^3\,dx = c\int_2^4 x^3\,dx$$

$$= c\,\frac{x^4}{4}\Big|_2^4 = c\left[\frac{4^4}{4} - \frac{2^4}{4}\right]^2$$

$$\Rightarrow 60c = 1 \Rightarrow c = 1/60$$

$P(X>3)$?

$$\int_3^4 \frac{1}{60}x^3\,dx = \frac{175}{240} \approx 0.729$$

Commulative dist function (cdf)

$X \longrightarrow$ CRV

cdf $\Rightarrow F(x) = P(X \le x) = \int_{-\infty}^{x} f(s)\,ds$

1. $P(X \le a) = F(a)$
2. $P(X > a) = 1 - F(a) = 1 - P(X \le a)$
3. $P(a \le X \le b) = \int_a^b f(x)\,dx = \int_{-\infty}^{b} f(x)\,dx - \int_{-\infty}^{a} f(x)\,dx = F(b) - F(a)$

$\uparrow f(x)$

Ex: T = time to failure of a system in hr

$$\boxed{f(t) = \mu \cdot e^{-\mu t}, \quad t \ge 0, \quad \mu > 0}$$

$$F(t) = P(T \le t) = \int_{-\infty}^{t} f(s)\,ds = \int_0^t \mu \cdot e^{-\mu s}\,ds$$

$$= \frac{\cancel{\mu} \cdot e^{-\mu s}}{-\cancel{\mu}}\Big|_0^t = 1 - e^{-\mu t}$$

$\Rightarrow$ cdf of exponential distribution

$$P(t_1 \le T \le t_2) = F(t_2) - F(t_1) = 1 - e^{-\mu t_2} - (1 - e^{-\mu t_1}) = \boxed{e^{-\mu t_1} - e^{-\mu t_2}}$$

Expected value and Varience

$X \to$ CRV $\quad \boxed{E[X] = \int_{-\infty}^{\infty} x \cdot f(x)\,dx}$

$Var(X) = \int_{-\infty}^{\infty} (x - E[X])^2 f(x)\,dx = \boxed{E[X^2] - E^2[X]}$

where $E[X^2] = \int_{-\infty}^{\infty} x^2 f(x)\,dx \qquad SD = \sigma_x = \sqrt{Var(X)}$

Ex: (cont.) $f(t) = \lambda e^{-\lambda t}, \; t \geq 0$ (T = time to failure)

$E[T] = \int_0^{\infty} t \cdot f(t)\,dt = \int_0^{\infty} t \cdot \lambda \cdot e^{-\lambda t}\,dt$

$$= \begin{cases} \int u\,dv = uv - \int v\,du \\ u = t \longrightarrow du = dt \\ dv = \lambda \cdot e^{-\lambda t}\,dt \longrightarrow v = -e^{-\lambda t} \end{cases}$$

$= t(-e^{-\lambda t}) \Big|_0^{\infty} - \int_0^{\infty} -e^{-\lambda t}\,dt = t e^{-\lambda t} \Big|_{\infty}^{0} + \int_0^{\infty} e^{-\lambda t}\,dt$

$= (0 - 0) + \frac{1}{\lambda} e^{-\lambda t} \Big|_{\infty}^{0} = \frac{1}{\lambda} \qquad E[T] = \frac{1}{\lambda}$

e.g., $f(t) = \frac{1}{2000} \cdot e^{-t/2000} \Rightarrow \lambda = \frac{1}{2000} \qquad E[T] = \frac{1}{\lambda} = 2000$ hrs

Uniform distribution

a CRV $X \sim \text{uni}(a,b)$ if the Pdf of X is as follows

$$f(x) = \frac{1}{b-a}, \quad a \le x \le b$$

$$F(x) = P(X \le x) = \int_a^x f(t)\,dt = \int_a^x \frac{1}{b-a}\,dt = \boxed{\frac{x-a}{b-a}}$$

$$P(x_1 \le X \le x_2) = F(x_2) - F(x_1) = \boxed{\frac{x_2 - x_1}{b-a}}$$

mean

$$\boxed{E[x] = \frac{a+b}{2}}$$

σ^2

$$\boxed{Var(X) = \frac{(b-a)^2}{12}}$$

Ex: Opening Price of a particular stock

$X \sim \text{uni}(35, 44)$

1. $P(X \le 40) = F(40) = \frac{x-a}{b-a} = \frac{40-35}{44-35} = 5/9$

2. $P(40 \le X \le 42) = \frac{42-40}{44-35} = 2/9$

3. $E[x] = \frac{a+b}{2} = \frac{35+44}{2} = 39.5$

4. $Var(X) = \frac{(b-a)^2}{12} = \frac{(44-35)^2}{12} = 6.75$

5. what Price is exceeded by 90% of the stocks?

Find x s.t. $P(X > x) = 0.9$

$$1 - F(x) = 0.9 \longrightarrow F(x) = 0.1$$

$$F(x) = \frac{x-a}{b-a} = \frac{x-35}{44-35} = 0.1 \longrightarrow x = 35.9$$

Exponential dist

A CRV $T \sim Exp(\lambda)$ if

$$\boxed{\text{Pdf } f(t) = \lambda e^{-\lambda t}, \quad t \geq 0, \quad \lambda > 0}$$

$$\text{cdf} = F(t) = P(T \leq t) = 1 - e^{-\lambda t}$$

$$E[x] = \frac{1}{\lambda} \qquad Var(T) = \frac{1}{\lambda^2}$$

Ex: Customer support center

Time between calls, $T \sim exp$ with a mean time between calls of 12 minutes

$E[T] = 12$ mins $\longrightarrow \lambda = \frac{1}{12}$ $\left(E[T] = \frac{1}{\lambda} \right)$

a) P(no calls arrive within 30 minutes) = ?

$$P(T > 30) = 1 - F(30) = 1 - (1 - e^{-\lambda t}) = e^{-\lambda t} = e^{-\frac{1}{12} \cdot 30}$$

$$= 0.082$$

b) P(at least one call within 10 min) = ?

$$P(T \leq 10) = F(10) = 1 - e^{-\lambda t} = 0.565$$

c) P(first call arrives between 5 and 10 minutes after opening) = ?

$$P(5 \leq T \leq 10) = F(10) - F(5) = 0.225$$

d) $E[X] = n \cdot p = n \cdot \frac{M}{N} = 15 \cdot \frac{30}{50} = 9$

$Var(X) = \left(\frac{N-n}{N-1}\right) \cdot np(1-p) = \frac{50-15}{49} \cdot 15 \cdot \frac{3}{5}\left(1-\frac{3}{5}\right)$

$= 2.5714 \Rightarrow \sigma_X = 1.6035$

$X + Y = 30$

$\boxed{Y = 30 - X}$

e) $E[Y] = E[30 - X] = E[30] - E[X] = 30 - 9 = 21$

$Var(Y) = V(30 - X) = 0 + V[-1 \cdot X] = (-1)^2 \cdot V(X) = V(X)$

$\sigma_Y = 1.6035$

Poisson dist

X = # of arrival for a given interval

$\mu = \lambda$: arrival rate (average # of arrivals) for that interval

$$\boxed{\begin{array}{l} \text{Pmf} : P(X = k) = \dfrac{e^{-\mu} \cdot \mu^k}{k!}, \quad k = 0, 1, 2, \dots \\ E[X] = \mu \quad Var(X) = \mu \end{array}}$$

in poisson dis $\mu = \lambda$

Ex: (cont.) $\mu = 10$ call / hr

a) P(5 calls in 1 hour) = ?

$\mu = 10/hr \Rightarrow P(X = 5) = \frac{e^{-10} \cdot 10^5}{5!} = 0.0378$

b) P(at most 3 calls during hour) = ?

$P(X \leq 3) = P(X=0) + P(X=1) + P(X=2) + P(X=3) = 0.0103$

c) P(5 calls in 30 minutes) = ?

$P(X = 5) = \frac{e^{-5} \cdot 5^5}{5!} = 0.1755$

$\mu = 10/hr = 5/30$ min

$\mu = 5$ avg

Ex: Cracks in a highway

$\mu = 2$ cracks/mile

a) P(no cracks in 5 miles) = ?

$\mu = 2/\text{mile} = 10/5 \text{ miles}$

$$P(X=0) = \frac{e^{-10} \cdot 10^{0}}{0!} = 4.5 \times 10^{-5}$$

$0! = 1$

b) P(at least one crack in 1/2 mile) = ?

$$P(X \geq 1) = 1 - P(X=0) = 1 - \frac{e^{-1} \cdot 1^{0}}{0!} = 0.632$$

$$1 - \frac{1}{e} = 0.632$$

c) E[cracks in 10 miles]

$\mu = 2/\text{mile} = 20/10 \text{ miles} = E[X] = \mu$

Poisson Aproximation of binomial

If a DRV is binomially distributed $X \sim bin(n, p)$ and

$n \longrightarrow$ large and $p \longrightarrow$ small then:

$$bin(x; n, p) \longrightarrow Poisson(x; \mu)$$

$$\boxed{\text{where } \mu = n \cdot p}$$

IMP on test

Ex: Let X be the # of rejects produced by a production line that makes 5,000 parts per day and 0.001 the Prob. that a part is unacceptable (p)

success

$$P(X=10) = \binom{5000}{10} (0.001)^{10} \cdot (0.999)^{5000-10} = 0.0181$$

$\mu = n \cdot p = 5000 \cdot 0.001 = 5$ $\qquad n = 5000$ (large) $p = 0.001$ (small)

$$P(X=10) = \frac{e^{-\mu} \cdot \mu^{x}}{x!} = \frac{e^{-5} \cdot 5^{10}}{10!} = 0.0181$$

Normal dist

Pdf

$$f(x) = \frac{1}{\sqrt{2\pi}\cdot\sigma} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad -\infty < x < \infty$$

$E[x] = \mu \quad Var(x) = \sigma^2$

Standard Normal variable

To Normalize $X \longrightarrow Z = \frac{X-\mu}{\sigma}$

The Standard Normal distribution

Let $\mu = 0$ and $\sigma^2 = 1 \longrightarrow Z \sim N(0,1)$ = Standard Normal

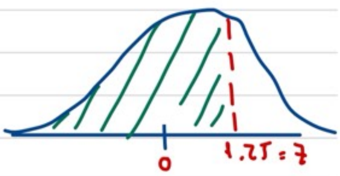
$$f(z) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{z^2}{2}}, \quad -\infty < z < \infty$$

The cdf of Z is denoted by $\Phi(z) = \phi = P(Z \le z)$

the values of $\phi(z)$ are tabulated; see book Appendix A.3

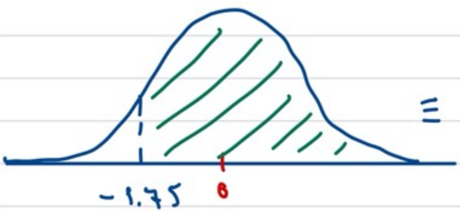
Ex:

a) $P(Z \le 1.25) = \phi(1.25) = 0.8944$

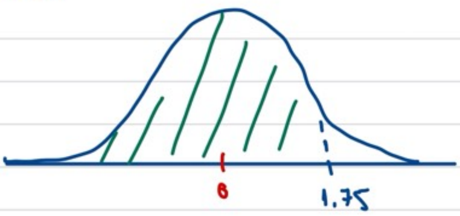


b) $P(Z > 1.25) = 1 - P(Z \le 1.25)$
$1 - \phi(1.25) = 0.1056$

c) $P(Z \ge -1.75) = P(Z \le 1.75)$



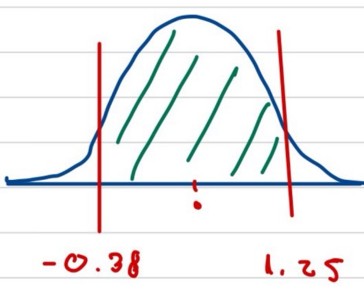
≡

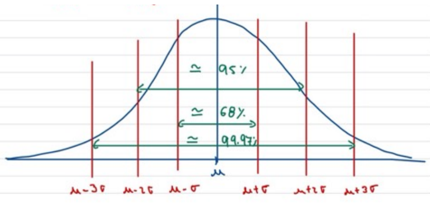


d) $P(Z \le -1.25) = P(Z > 1.25) = 0.1056$

e) $P(-0.38 \le Z \le 1.25) = \phi(1.25) - \phi(-0.38)$

$= 0.8944 - 0.3520 = 0.5424$ ↗





⟶

$P(\mu-\sigma \le x \le \mu+\sigma) = P(-\sigma \le x-\mu \le \sigma) = P(-1 \le \frac{x-\mu}{\sigma} \le 1)$

$P(-1 \le Z \le 1) = 2\,P(0 \le Z \le 1) = 2\,[P(Z \le 1) - P(Z \le 0)]$

$2\,[0.8413 - 0.5] = 0.6826 \simeq 68\%$

Non-Standard Normal Distribution

$X \sim N(\mu, \sigma^2) \Rightarrow E[X] = \mu \quad var(X) = \sigma^2$

$P(a \le X \le b) = ?$

1. Let $Z = \frac{X-\mu}{\sigma} \longrightarrow E[Z] = E\left[\frac{X-\mu}{\sigma}\right] = \frac{1}{\sigma} E[X-\mu]$

$= \frac{1}{\sigma}(E[X] - \mu) = \frac{1}{\sigma}(\mu - \mu) = 0$

$Var(Z) = Var\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma^2}\left(Var(X) - Var(\mu)\right)$

$= \frac{1}{\sigma^2} \cdot (\sigma^2 - 0) = 1$

$\boxed{\text{Then } Z \sim N(0,1)}$

2. $P(a \le X \le b) = P\left(\frac{a-\mu}{\sigma} \le \frac{X-\mu}{\sigma} \le \frac{b-\mu}{\sigma}\right)$

$= P\left(\frac{a-\mu}{\sigma} \le Z \le \frac{b-\mu}{\sigma}\right)$

$$\boxed{= \phi\left(\frac{b-\mu}{\sigma}\right) - \phi\left(\frac{a-\mu}{\sigma}\right)}$$

Ex: time to react to the break lights on a decelerating vehicle, sec

$X \sim N(1.25, 0.46^2)$
σ

$E[X] = \text{mean} = 1.25 \qquad SD(X) = 0.46 \text{ sec}$

$P(\overset{b}{1.00} \le X \le \overset{a}{1.75}) = P\left(\frac{\overset{b-\mu}{1-1.25}}{\underset{\sigma}{0.46}} \le Z \le \frac{\overset{a-\mu}{1.75-1.25}}{\underset{\sigma}{0.46}}\right)$

$= P(-0.54 \le Z \le 1.09)$

$= \phi(1.09) - \phi(-0.54) = 0.8621 - 0.2946$

$= 0.5675$

L7?

Recall:

Negative binomial RV: X = # of failures that precede the r^{th} success

$$\text{pmf: } P(X=x) = \binom{x+r-1}{r-1} p^r (1-p)^x \quad , \; x = 0,1,2,\ldots$$

$$E[X] = \frac{r(1-p)}{p} \qquad Var(X) = \frac{r(1-p)}{p^2}$$

Hypergeometric RV: X = # of type 1 in a random sample of size "n"

$$\text{pmf: } P(X=x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}} \qquad E[X] = n \cdot p = n \cdot \left(\frac{M}{N}\right)$$

$$Var(X) = \left(\frac{N-n}{N-1}\right) \cdot n \cdot p \cdot (1-p)$$

Lecture 8

Recall:

Negative binomial RV: X = # of failures that precede the r^{th} success

$$\text{pmf: } P(X=x) = \binom{x+r-1}{r-1} p^r (1-p)^x \quad , \; x = 0,1,2,\ldots$$

$$E[X] = \frac{r(1-p)}{p} \qquad Var(X) = \frac{r(1-p)}{p^2}$$

Hypergeometric RV: X = # of type 1 in a random sample of size "n"

$$\text{pmf: } P(X=x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}} \qquad E[X] = n \cdot p = n \cdot \left(\frac{M}{N}\right)$$

$$Var(X) = \left(\frac{N-n}{N-1}\right) \cdot n \cdot p \cdot (1-p)$$

Recall L9 Lecture 9

DRV: Poisson distribution $\longrightarrow$ arrival in a time interval

$E[X] = \mu = Var(X) \quad P(X=k) = \dfrac{e^{-\mu} \cdot \mu^{k}}{k!}, \quad k = 0, 1, 2, \dots$

Poisson approximation to binomial

$n \to$ large $\quad p \to$ small $\quad \therefore \mu = n \cdot p$

CRV: $P(a \le X \le b) = \int_a^b f(x)\,dx, \quad P(X=k) = 0$

$$P(X \le x) = \int_{-\infty}^{x} f(s)\,ds = F(x)$$

$$P(X > a) = 1 - F(a)$$

$$P(a \le X \le b) = \int_{-\infty}^{b} f(x)\,dx - \int_{-\infty}^{a} f(x)\,dx = F(b) - F(a)$$

$$E[X] = \int_{-\infty}^{\infty} x\,f(x)\,dx \qquad V(X) = \int_{-\infty}^{\infty} (x - E[X])^2 f(x)\,dx = E[X^2] - E^2[X]$$

$$E[X^2] = \int_{-\infty}^{\infty} x^2 f(x)\,dx \qquad \boxed{\int u\,dv = uv - \int v\,du}$$

Uniform distribution

$X \sim \text{uni}(a, b) \quad f(x) = \dfrac{1}{b-a}, \quad a \le x \le b$

$F(x) = \dfrac{x-a}{b-a} \qquad P(x_1 \le X \le x_2) = \dfrac{x_2 - x_1}{b-a} \qquad E[X] = \dfrac{a+b}{2}$

$$\sigma_x^2 = \frac{(b-a)^2}{12}$$

Exponential Distribution

$f(t) = \lambda e^{-\lambda t}, \quad t \ge 0, \quad \lambda > 0$ $\quad$ * lifetime, time between events

cdf $\longrightarrow F(t) = 1 - e^{-\lambda t} \qquad E[t] = \dfrac{1}{\lambda} \qquad Var(t) = \dfrac{1}{\lambda^2}$