Poisson Distribution.

$X \sim \text{Poiss}(\lambda)$ $\qquad$ λ "rate parameter"

$\lambda > 0$

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

Is this a valid PMF ?

$$P(X = k) \geq 0$$

$$\sum_{k=0}^{\infty} \frac{e^{-\lambda} \lambda^k}{k!} = e^{-\lambda} \underbrace{\sum_{k=0}^{\infty} \frac{\lambda^k}{k!}}_{\text{Taylor series for } e^{\lambda}} = e^{-\lambda} e^{\lambda} = 1.$$

$$E[X] = \sum_{k=0}^{\infty} k \frac{e^{-\lambda} \lambda^k}{k!}$$

$$= e^{-\lambda} \sum_{k=0}^{\infty} \frac{k \lambda^k}{k!}$$

$$= e^{-\lambda} \sum_{k=1}^{\infty} \frac{k \lambda^k}{k!}$$

$$= e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!}$$

$$= \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!}$$

$$j = k - 1$$

$$\lambda e^{-\lambda} \sum_{j=0}^{\infty} \frac{\lambda^j}{j!}$$

$$\lambda e^{-\lambda} e^{\lambda} = \lambda$$

$$\underline{E[X] = \lambda.}$$

Why is the Poisson Useful?

- applications where we're counting the number of "successes", where there are a large number of trials, and each trial has a small probability of success.

Examples.

- # injuries / year in a factory.
- # emails received in an hour
- # memory errors on a memory chip in a spacecraft.
- # chocolate chips in a cookie

The Poisson distribution is a **model**.

this is an important concept when applying probability to the real world.

Consider a set of events

$A_1, A_2 \ldots A_n$ $\qquad P(A_j) = p_j$

n large

p_j small

Events are independent or "weakly dependent"

of the A_j's that occur is approximately

Poiss (λ)

$$\lambda = \sum_{j=1}^{n} p_j$$

If the events are independent, and $p_j = p$ for all j

$X \sim \text{Bin}(n, p)$

let $n \to \infty$

$p \to 0$

keeping np constant.

call $\lambda = np$

$$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

$$= \frac{n(n-1)(n-2)\ldots(n-k+1)}{k!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k}.$$

take limit $n \to \infty$

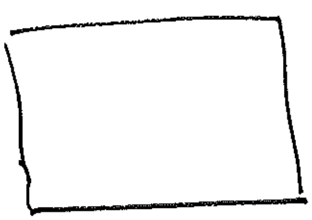
$$\frac{n^k}{k!}$$

$$\left(1-\frac{\lambda}{n}\right)^{n-k} = \underbrace{\left(1-\frac{\lambda}{n}\right)^{n}}_{\to e^{-\lambda}} \underbrace{\left(1-\frac{\lambda}{n}\right)^{-k}}_{\to 1}.$$

$$p(x=k) = \frac{n^k}{k!}\frac{\lambda^k}{n^k}e^{-\lambda}$$

$$= \frac{\lambda^k}{k!}e^{-\lambda}$$

$\to$ Poisson distribution is limit of Binomial as $n \to \infty$, $p \to 0$ in such a way that the expected number of successes stays fixed $(=\lambda)$



flux of 10^{10} particles on the chip

prob. of a bit flip from 1 particle is $\sim 10^{-12}$

$$\mathrm{Bin}(10^{10}, 10^{-12}) \simeq \underbrace{\mathrm{Poisson}(10^{-2})}_{\text{much easier to compute.}}$$

Continuous Distributions.

Discrete PMF $P(X = x)$

replace this with a Probability Density Function PDF

CDF $F(x) = P(X \leq x)$ — also valid for a continuous RV

Definition of PDF.

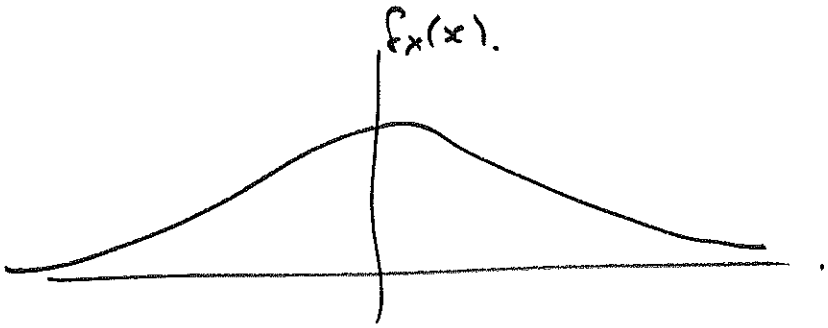
RV X has PDF $f_X(x)$.

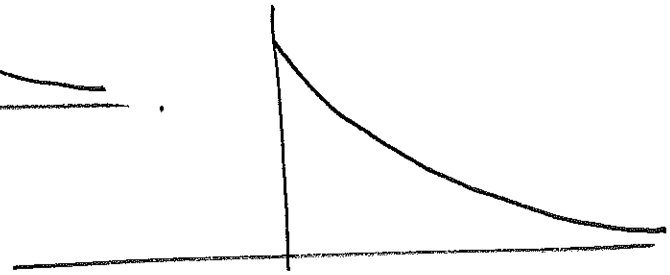
if $P(a \leq X \leq b) = \int_a^b f_X(x)\,dx$ for all a, b

for continuous RVs, probabilities are given by integrals.

valid ~~PDF~~ PDF: $f_X(x) \geq 0$

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

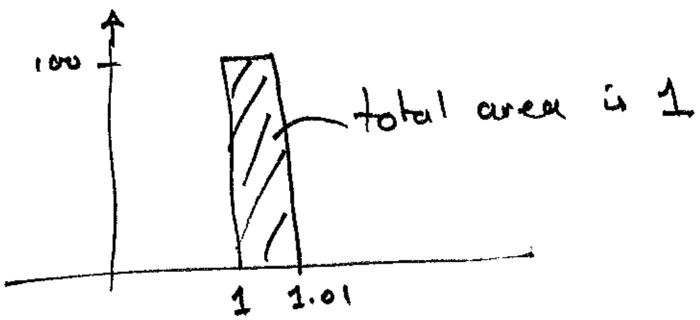


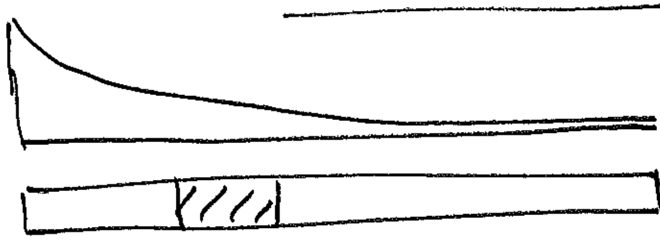


area underneath the curve = 1

$f_X(x)$ may be > 1 for some values of x.

Probability density.



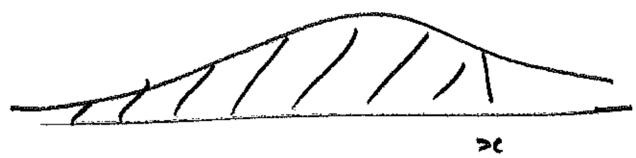


non-uniform density

X has PDF $f_X(x)$.

the CDF $F(x) = P(X \leq x) = \cancel{\int_{-\infty}^{x} f_X(x)}\, dx.$

$$= \int_{-\infty}^{x} f_X(t)\, dt$$



<u>Converse.</u>

$$f_X(x) = F'(x)$$

$$P(a < x < b) = \int_a^b f_X(x)\, dx$$

$$= F(b) - F(a).$$

$$E[X] = \sum x\, P(X = x)$$

$$E[X] = \int_{-\infty}^{\infty} x\, f_X(x)\, dx$$

Variance - measure of spread.

$$E[X - E[X]] = 0$$

$$E[X] - E[\underbrace{E[X]}_{\text{constant}}]$$

$E[\text{constant}]$ = the constant value.

$$E[X] - E[X] = 0$$

$$E[(X - E[X])^2] = Var(X)$$ this is in (units of $X)^2$

$$STD(X) = \sqrt{Var(X)}$$ - this is in the same units as X.

$$E[X^2 - 2X\,E[X] + (E[X])^2]$$

apply linearity of expectation

$$E[X^2] - 2\,E[X]\,E[X] + (E[X])^2.$$

$$Var(X) = E[X^2] - (E[X])^2$$

now $var(X) \geq 0$ hence $E[X^2] \geq (E[X])^2.$

Continuous Uniform Distribution.

Unif(a, b)

How do we define "completely random in the interval $[a, b]$"

when $P(X = x) = 0$ for all points in the interval?

→ Probability that X takes a value in an interval is proportional to the length of that interval.

$$f_X(x) = \begin{cases} c & a \le x \le b \\ 0 & \text{otherwise.} \end{cases}$$

for $f_X(x)$ to be a valid pdf

$$1 = \int_a^b c \, dx = c(b-a)$$

$$\Rightarrow c = \frac{1}{b-a} = \frac{1}{\text{length of the interval}}$$

if $a = 0$, $b = 1$.

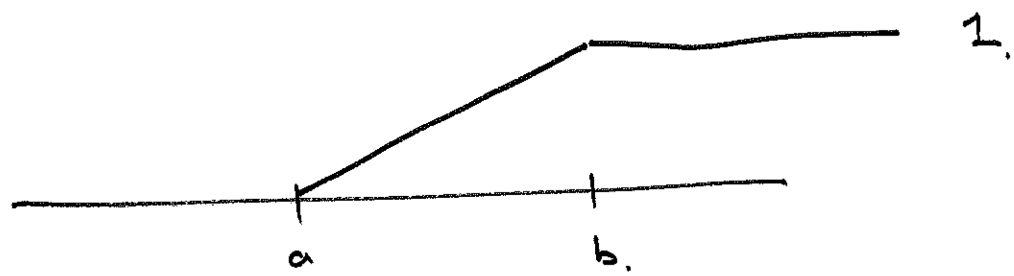
then $f_X(x) = 1 \quad 0 \le x \le 1$

$= 0 \quad$ otherwise.

CDF $\quad F(x) = \int_{-\infty}^{x} f(t)\, dt.$

$$= \int_{a}^{x} f(t)\, dt \qquad \text{for } a \le x \le b$$

$$F(x) = \begin{cases} 0 & x < a. \\ \dfrac{x-a}{b-a} & x \le a \le b. \\ 1 & x > b. \end{cases}$$



$$E[x] = \int_{-\infty}^{\infty} x\, f_x(x)\, dx$$

$$= \int_{a}^{b} \frac{x}{b-a}\, dx$$

$$= \frac{1}{b-a} \left. \frac{x^2}{2} \right|_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{(b-a)(b+a)}{2(b-a)}.$$

$$= \frac{b+a}{2} \qquad \text{midpoint of the interval}$$

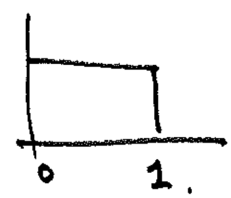
$$\mathrm{Var}[X] = E[X^2] - (E[X])^2$$

$$E[X^2] = \int_{-\infty}^{\infty} x^2 \times \frac{1}{b-a)} \, dx$$

$$= \int_a^b \frac{x^2}{b-a} \, dx$$

$$= \frac{1}{b-a} \left. \frac{x^3}{3} \right|_a^b .$$

$$= \frac{b^3 - a^3}{3(b-a)}.$$



specific case of $a = 0$, $b = 1$

$$E[X^2] = \frac{1}{3}$$

$$\mathrm{Var}[X] = \frac{1}{3} - \left(\frac{1}{2}\right)^2 = \frac{1}{12}.$$

$$E[X^2] = \int x^2 f_X(x) \, dx$$

$$E[g(X)] = \int g(x) f_X(x) \, dx.$$

← this is always true.

(will prove it later).

If we can generate RVs that follow the Unif(0,1) distribution, we can use them to generate RVs that follow any distribution (in principle)

$U \sim \text{Unif}(0,1)$

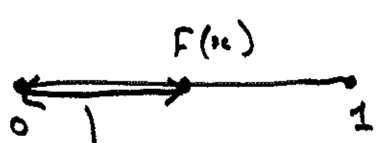
F - the CDF of the distribution we're interested in.

let $x\cancel{X} = F^{-1}(u)$ ← generate realizations, u, from Unif(0,1). Plug them in to $F^{-1}()$

what comes out are realizations of a RV with distribution F.

Proof.

$$P(X \le x) = P(F^{-1}(U) \le x)$$
$$= P(U \le F(x))$$ apply $F()$ to both sides.



Prob of a unif(0,1) being in the region $[0, F(x)]$ is $F(x)$.

hence $P(X \leq x) = F(x)$

- useful if we can find $F^{-1}()$ analytically.

- ~~[illegible]~~

realizations of non-uniform RV. are very useful for numerically approximating expectations

$$E[x] = \int x f(x)\, dx$$

$$E[g(x)] = \int g(x) f(x)\, dx$$

$$\simeq \frac{1}{n} \sum g(x_i)$$ where x_i are ~~[illegible]~~ realizations of $X \sim f(x)$.

Monte Carlo Method.