TAM 598 Lecture 4 :

CONTINUOUS RANDOM VARIABLES

Announcements:

- HW 1 covers lectures 1-4 ; due on Feb 12
- to be submitted via CANVAS

I. CONTINUOUS RANDOM VARIABLES

Let X be a random variable that can take values in $\mathbb{R}$. If the range of X is uncountable (e.g. if it forms an interval), we say X is a continuous random variable

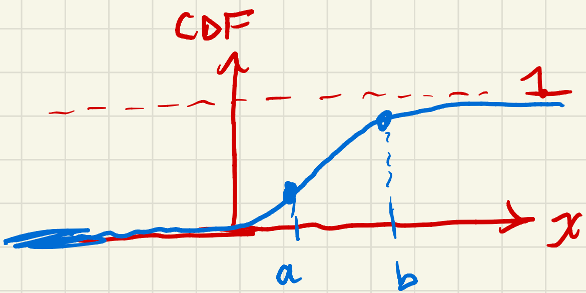
eg X = mass of a ball bearing, the temperature of a room, ...

1) cumulative distribution function (CDF) of random variable X is denoted by $F_X(x)$ and is the probability that X takes a value less than or equal to x

$F(x)$ or $F_X(x)$

uppercase lowercase

$$F_X(x) = p(X \le x) = p(\{w : X(w) \le x\})$$



properties of CDF:

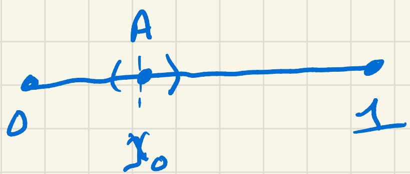
- $F_X(x)$ is an increasing function
- $F_X(x \to -\infty) = 0$

  $F_X(x \to +\infty) = 1$
- $p(a \le X \le b) = F_X(b) - F_X(a)$

2) probability density function ← PDF is a function $f_X(x)$ that gives us the probability that X lies in any Borel subset A of $\mathbb{R}$

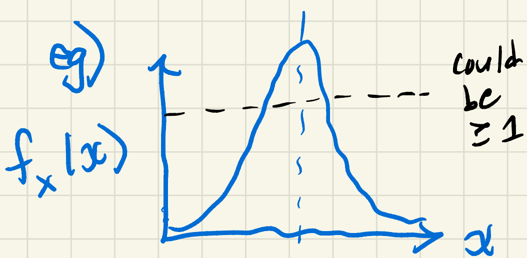
$p(x)$

generated by open sets



$$P(X \in A) = \int_A f_X(x)\,dx$$

eg)



properties of pdf:

1. $p(x) \geq 0$ for all x
2. $\displaystyle\int_{-\infty}^{\infty} p(x)\,dx = 1$
3. $\boxed{\dfrac{d}{dx} F_X(x) = p(x)}$ (*)

Fund Theorem of Calculus

$$\int_a^b f'(t)\,dt = f(b) - f(a)$$

know: $F(x) = \int_{-\infty}^{x} f_x(t)\,dt$

take $\frac{d}{dx}$: $F'(x) = \frac{d}{dx}\int_{-\infty}^{x} f_x(t)\,dt \xrightarrow{\text{FTOC}} = f_x(x)$

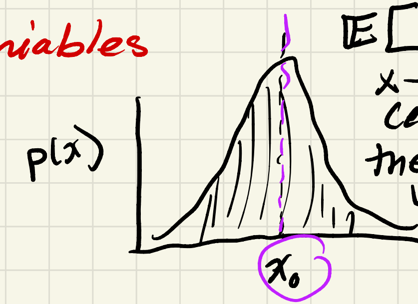
3) Expectations of Continuous Random Variables

recall discrete RV:

$$\mathbb{E}[X] = \sum_x x\, p(x)$$

here:

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



$\mathbb{E}[X] =$ x-coord of centroid of the area under curve

again, $\mathbb{E}[X]$ need not be a value that X is allowed to take.

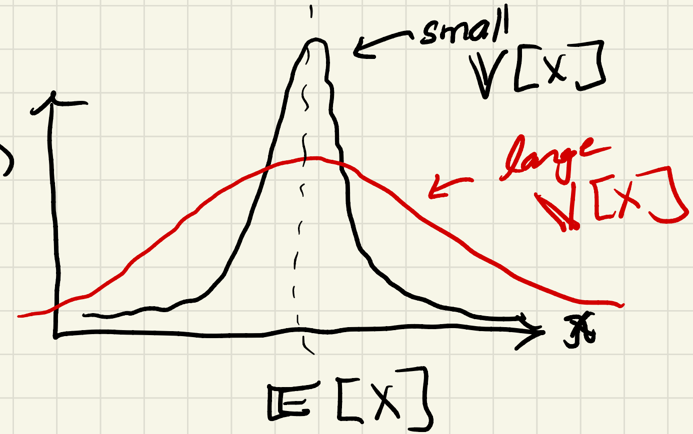
also: recall discrete RV:

$$\mathbb{E}[g(X)] = \sum_x g(x)\, p(x)$$

here:

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

4) Variance of continuous random variables is again the average squared distance of X from its expected value. Tells us how "spread out" X is



$$V[X] = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right]$$

$$= \int_{-\infty}^{\infty} (x - \mathbb{E}[X])^2 \, p(x)\, dx$$

⊙ units of variance? X^2

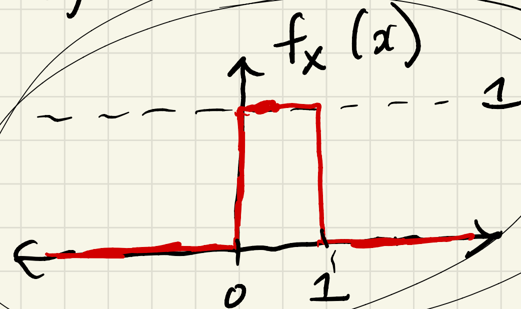
⊙ we define $\sigma_X = \sqrt{V[X]}$ as the standard deviation of X. σ_X has the same units as X.

II. THE UNIFORM DISTRIBUTION – a random variable equally likely to take a value within a given interval.

Between 0 and 1 :

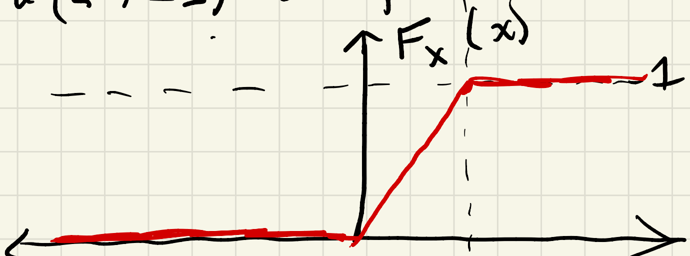
⊙ $X \sim \mathcal{U}([0,1])$ has pdf $= \begin{cases} 1 & 0 \le x \le 1 \\ 0 & \text{else} \end{cases}$

"$\sim$" "is distributed as"
"is drawn from"

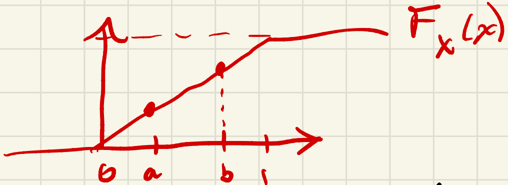


⊙ $X \sim \mathcal{U}([0,1])$ has cdf

$$F(x) = \int_0^x f_X(u)\, du$$
$$= \int_0^x du = x$$



⊙ The probability that X takes values in $[a, b]$ for $a < b$ in $[0, 1]$ is



$$P(a \le x \le b) = F(b) - F(a) = b - a$$

⊙ The expectation is

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f_X(x)\, dx = \int_0^1 x\, dx = \frac{1}{2} x^2 \Big|_{x=0}^{x=1} = \frac{1}{2}$$

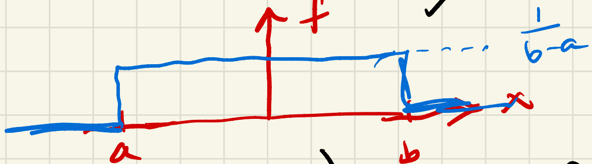
⊙ The variance is

$$\mathbb{V}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$$

$$= \frac{1}{3} - \left(\frac{1}{2}\right)^2 = \frac{1}{12}$$

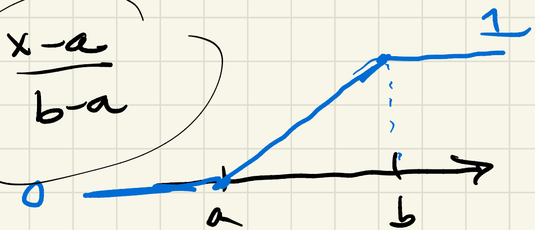
Between a and b

- $X \sim \mathcal{U}([a,b])$ has pdf $f_X(x) = \begin{cases} \frac{1}{(b-a)} & x \in [a,b] \\ 0 & \text{else} \end{cases}$



- $X \sim \mathcal{U}([a,b])$ has cdf

$$F(x) = \int_{-\infty}^{x} f_X(u)\,du = \int_{a}^{x} \frac{1}{b-a}\,du = \frac{x-a}{b-a}$$



- $$\mathbb{E}[X] = \int_a^b x\left(\frac{1}{b-a}\right)dx = \frac{1}{2}(a+b)$$

- $$\mathbb{V}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \frac{1}{12}(b-a)^2$$

Say we know how to obtain samples from $\mathcal{U}([0,1])$, but we want samples from $\mathcal{U}([a,b])$. How do we do it?

construct a map

let $Z \sim \mathcal{U}([0,1])$

define $X = \underline{a + (b-a)Z}$

sample Z, scale it according to X

pf: it is sufficient to show that $X = a + (b-a)Z$ has the same CDF as $\mathcal{U}([a,b])$

$$F(x) = P(X \le x) = P(a + (b-a)Z \le x)$$

$$= P((b-a)Z \le x - a)$$

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

III. THE NORMAL DISTRIBUTION a.k.a. THE GAUSSIAN DISTRIBUTION

$$\boxed{X \sim N(\mu, \sigma^2)}$$

μ = expectation
σ^2 = variance
be aware: ~~$N(\mu, \sigma)$~~ sipy

- appears over and over again in nature, in measurements, ...
- why?

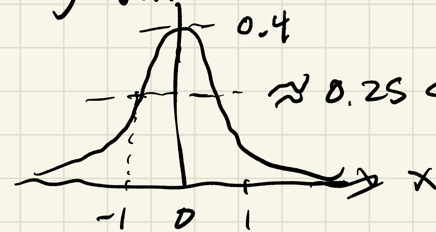
1) say you know only a distribution's expectation and variance. Amongst all possible governing distributions, the gaussian dist is the one that maximizes the uncertainty (ie assumes no extra structure, is the "least biased")

2) when summing many independent variables, the result is normal (least structed)

CENTRAL LIMIT THEOREM

standard normal distribution

$Z \sim N(0,1)$ mean $\mu = 0$, variance $\sigma^2 = 1$



PDF $\phi(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right)$

CDF $\Phi(z) = P(Z \le z) = \int_{-\infty}^{z} \phi(z')\,dz'$ no closed form solution

general normal distribution

$X \sim N(\mu, \sigma^2)$ mean μ, variance σ^2

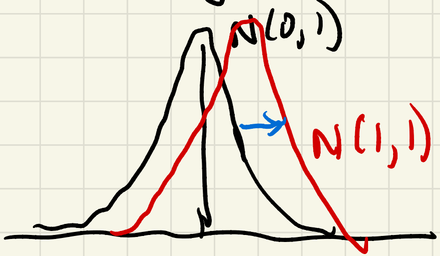
PDF $f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$

→ consider the pdf of $N(0,1)$:

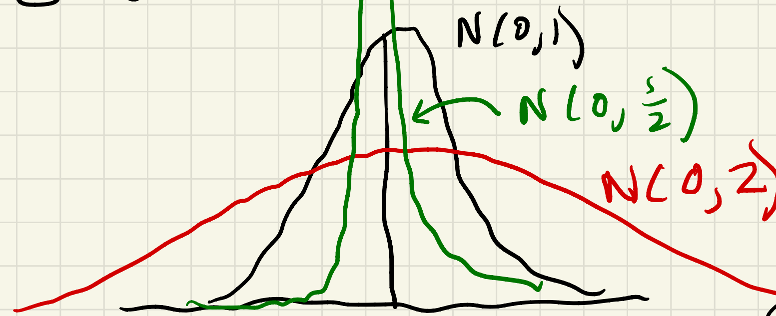
$$\phi(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right)$$

$$\frac{1}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right) = \underbrace{\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)}_{f_X(x) \text{ of } N(\mu,\sigma^2)}$$

so: ① μ shifts distribution right or left



② σ^2 scales the distribution



Quantiles of the Normal Distribution

Find the value $Z = z_q$ so that $P(Z \le z_q) = q$, ie

find $\Phi(z_q) = q$

z_q is called the **q-quantile**

want $z_q = \Phi^{-1}(q)$

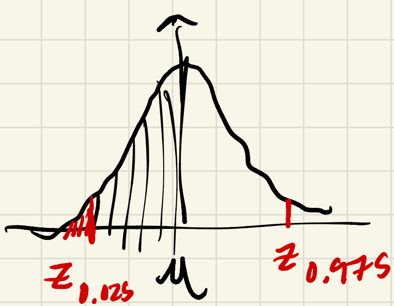
eg) $Z_{0.50}$ is called the median (and here coincides w/ the expectation)

eg) $Z_{0.025}, Z_{0.975}$ is the 95% confidence interval

$$P(Z_{0.025} \le Z \le Z_{0.975}) = \Phi(Z_{0.975}) - \Phi(Z_{0.025})$$

$$= 0.975 - 0.025$$

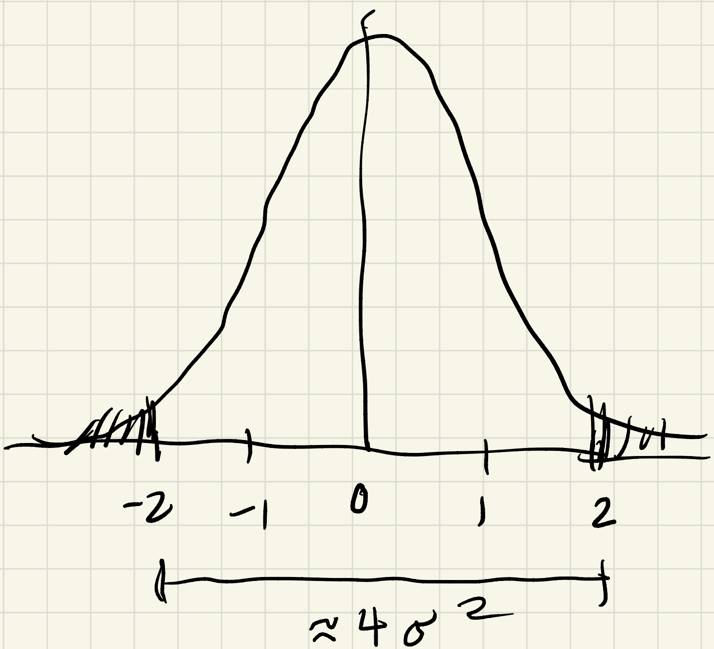
$$= 0.95$$



For $\underline{N(0,1)}$

$z_{0.975} \approx 1.96$

$z_{0.025} \approx -1.96$



Getting any normal from the standard normal:

let $Z \sim N(0,1)$

define $X = \mu + \sigma Z$

change of variables

Then: $X \sim N(\mu, \sigma^2)$

Pf: consider the CDF of X

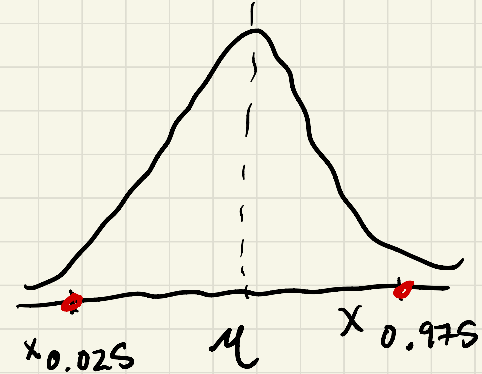
$$F_X(x) = P(X \le x) = P(\mu + \sigma Z \le x)$$

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

take the derivative:

$$f_X(x) = F_X'(x) = \frac{1}{\sigma}\,\phi\left(\frac{x-\mu}{\sigma}\right)$$

How to find a q-quantile X_q of X:



$$P(X \le x_q) = q$$

$$P(\mu + \sigma Z \le x_q) = q$$

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

$$\Phi\left(\frac{x_q - \mu}{\sigma}\right) = q \implies$$

$z_q = \dfrac{x_q - \mu}{\sigma}$ is the q-quantile of Z

$$x_q = \mu + \sigma z_q$$