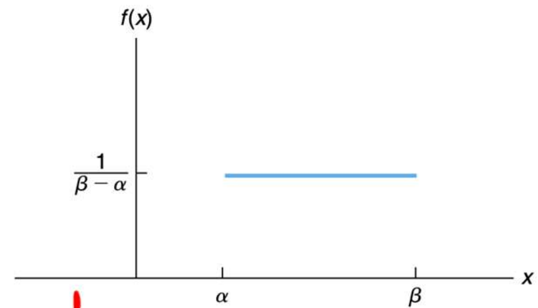
Special Continuous Random Variables

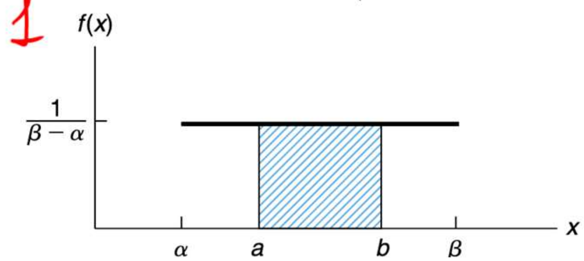
Uniform Random Variable

- X is uniformly distributed between α and β
- $X \sim U(\alpha, \beta)$
- $P(X) = \frac{1}{\beta-\alpha}$

$$\int_x p(x)\,dx = 1$$

$$= \int_{x=\alpha}^{x=\beta} p(x)\,dx = \int_{\alpha}^{\beta} \frac{1}{\beta-\alpha}\,dx = 1$$





- $P(X \in [a, b]) = \frac{b-a}{\beta-\alpha}$

- E[X] =

$$\int_x x\,p(x) = \int_{x=\alpha}^{\beta} x \frac{1}{\beta-\alpha}\,dx = \frac{x^2}{(\beta-\alpha)2}\Big]_{\alpha}^{\beta} = \frac{\beta^2-\alpha^2}{(\beta-\alpha)2} = \frac{\alpha+\beta}{2}$$

Variance of uniform random variable

$$E(X^2) = \int_{x=\alpha}^{\beta} x^2 \left(\frac{1}{\beta-\alpha}\right) dx = \frac{1}{\beta-\alpha}\left(\frac{x^3}{3}\right)\Bigg|_{\alpha}^{\beta} = \frac{\beta^3-\alpha^3}{(\beta-\alpha)3}$$

$$= \frac{\beta^2+\alpha^2+\alpha\beta}{3}$$

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

$$= \frac{\beta^2+\alpha^2+\alpha\beta}{3} - \left(\frac{\alpha+\beta}{2}\right)^2 = \frac{(\alpha-\beta)^2}{12}$$

Doubt

An example application of uniform R.V.s

- Given a set n elements $x_1, x_2, \ldots, x_n$. You need to write an algorithm for selecting a random subset k of the n elements given access to a uniform random number generator U(0,1)
- R = ϕ

for $i = 0$ to n

$u_i \sim U(0,1)$

$r_i = |R|$

$p_i = \frac{k - r_i}{n - i}$

if $u_i < p_i$ add x_i to R

stop if $|R| = k$

$\binom{n}{k} = \frac{n!}{(n-k)!\,k!} = N$

All possible subset of size k

- $S_0, S_1, \ldots, S_{N-1}, S_N$
- $u \sim U(0,1)$
- If $u \in \left[\frac{j}{N}, \frac{j+1}{N}\right]$ then choose set S_j

Selecting a random subset in a streaming setting

Say you are hosting a webserver. You want to track the interaction of a random subset k of customers that arrive at the webserver. But you do not know the number of customers that will arrive in advance.

You have limited memory k and cannot store all possible customers data that arrive and then select a subset.

You can generate uniform random numbers between 0 and 1.

$R \leftarrow \{x_1, x_2 - - x_k\}$

for $i = k+1$ to ∞

choose a random # between $1 - - k+1$

reject customer from $[R\ x_i]$

Reservoir sampling: n is unknown, data arrives in a stream

$R_k \leftarrow \{x_1, x_2 - \quad x_k\}$

Let R_i denote the state of reservoir R after seeing x_i

- Initialize R_k with first k elements.
- Foreach subsequent x_i
  - Sample a uniform integer s from 1,2,..i.
  - If $s \leq k, R[s] = x_i$

Prove that the probability with which we add element x_j to R_i after seeing i elements is $\frac{k}{i}$

By induction:

- Base case $i = k$ holds
- ~~Assume~~ that at $i-1$ $P(x_j \in R_{i-1}) = \frac{k}{i-1}$ $j \in [1..i-1]$

We assumed to insert it at the i-1th element seeing.

- $P(x_j \in R_i) = P(x_j \in R_{i-1}) \cdot \left(1 - \frac{k}{i} \cdot \frac{1}{k}\right)$

Trying to calculate if we add it into the ith seeing then we'd get the formula required.

A big doubt, how we did it! The induction step.

$$= \frac{k}{i-1}\left(1-\frac{1}{i}\right)$$

$$= \frac{k}{i}$$

Normal (Gaussian) Random Variable

Normal distribution.

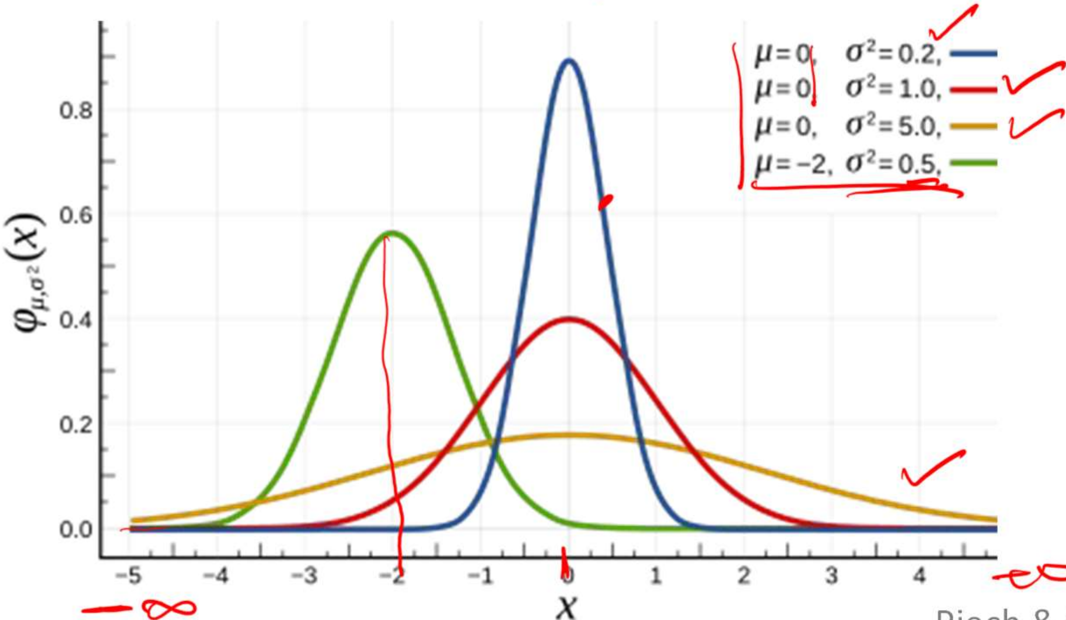
mean

variance

Support:
$(-\infty, \infty)$

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

density



$-\infty$ ∞

Normal (Gaussian) Random Variable

Support:
$(-\infty, \infty)$

mean
variance

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

PDF:

$$f(X = x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

Anatomy of a The Normal PDF

$$f(X = x) = \underbrace{\frac{1}{\sigma\sqrt{2\pi}}}_{\substack{\text{a constant:}\\ \text{makes the integral}\\ \text{over all possible}\\ \text{outcomes sum to 1}}} e^{\frac{-\overbrace{(x-\mu)^2}^{\substack{\text{distance to the mean}\\ \text{(makes the PDF symmetric}\\ \text{around the mean)}}}}{2\sigma^2}}$$

...normalized by the variance

Expected value of a normal distribution

verify that μ is the expected value of

$x \sim N(\mu, \sigma^2)$

$E([x-\mu]) = E(x) - \mu$

$$\int_{-\infty}^{\infty} (x-\mu) \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \left. \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma\,(\sigma^2)} \right|_{-\infty}^{+\infty} = 0$$

We did the tdt thing here!

$E[(x-\mu)] = 0 \Rightarrow E(x) = \mu$

Revisit Needed

I guess learning integral again is very much needed now!

Variance

$$E((X-\mu)^2) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x-\mu)^2 e^{-(x-\mu)^2/(2\sigma^2)} dx$$

$$= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y^2 e^{-(y)^2/(2)} dy = \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (y)(ye^{-(y)^2/(2)}) dy$$

u

v du

$$= \frac{\sigma^2}{\sqrt{2\pi}} \left[\left(-ye^{-y^2/2} \right)_{-\infty}^{\infty} - \int_{-\infty}^{\infty} -e^{-y^2/2} dy \right]$$

$$\int u\,dv = uv - \int v\,du$$

$$\int ye^{-y^2/2} dy = -e^{-y^2/2}$$

$$= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-y^2/2} dy = \frac{\sigma^2}{\sqrt{2\pi}} \sqrt{2\pi} = \sigma^2$$