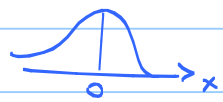
Lecture 4 (Ch. 1)

More examples of dists:

$\boxed{x = \text{cont. Easier}}$

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \quad -\infty < x < \infty \qquad \int f\,dx = 1 \checkmark \quad f \geq 0 \checkmark$$



Called $\boxed{\text{Standard normal distr.}}$

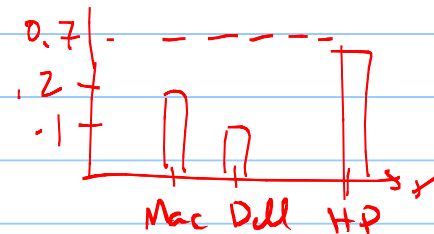
$\boxed{x = \text{categorical. Harder}}$

E.g. x = computer Brand

x	Mac	Dell	HP
$p(x)$	0.2	0.1	0.7

$p(x) \geq 0$ ✓

$\sum_x p(x) = 1$ ✓



Table

Note: There is no data anywhere here. These are not histograms

Chart

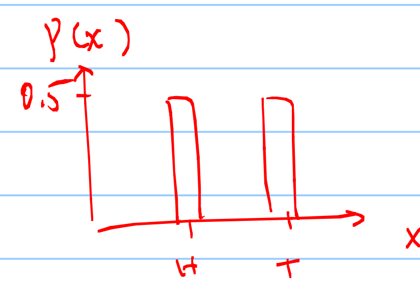
or formula

E.g. x = "state of a fair coin"

x	H	T
$p(x)$	$1/2$	$1/2$

$\boxed{\text{Bernoulli distr.}}$

Later, we will replace $1/2$ with something else.



E.g. x = "number of heads out of n tosses of a fair coin."

$$p(x) = \frac{n!}{x!(n-x)!} \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{n-x}$$

$\boxed{\text{Binomial distr.}}$

Later, we will replace the $\frac{1}{2}$ with other values.

We will derive this $p(x)$, later.

Don't forget; all these $p(x)$'s are used to describe the population of x.

Bernoulli and Binomial dists are important, but there are a few (more) special distributions which arise frequently either because they have desirable mathematical properties, or because there are lots of data in the real world whose histograms look like these distributions.

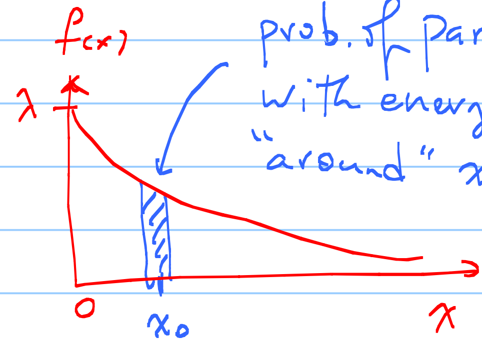
1) <u>Exponential</u> (family), x = cont.

E.g. Radiated Heat (ie. energy of particles) or, (inter-)arrival time between indep. events.

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{else} \end{cases}$$

Note the parameter: $\lambda > 0$

Meaning: $\lambda = \frac{1}{\text{mean } x}$ (later).



prob. of particles with energy "around" x_0.

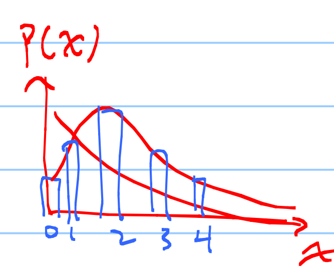
e.g. energy of particles coming out of uranium.

2) <u>Poisson</u>, x = discrete

e.g. # of bombs dropped over London per block.

$$\underbrace{p(x)}_{} = \frac{e^{-\lambda} \lambda^{x}}{x!}, \quad x = 0, 1, 2, \ldots$$

prob. of x bombs dropped per block.



parameter: $\lambda > 0$

meaning: λ = mean x (later)

3) **Binomial (revisited)** $x =$ discrete

We'll derive its mass function next time, but it's :

E.g. # of Heads out of n tosses.

# of defective gates on a chip with n gates

# of girls in a sample of size n.

Etc.

$$\underbrace{p(x)}_{\text{prob of } x \text{ heads out of } n \text{ tosses.}} = \boxed{\frac{n!}{x!\,(n-x)!}} \cdot \pi^{x} (1-\pi)^{n-x} \quad , \quad x = 0, 1, \cdots n$$

# of Hs out of n tosses

$\binom{n}{x}$

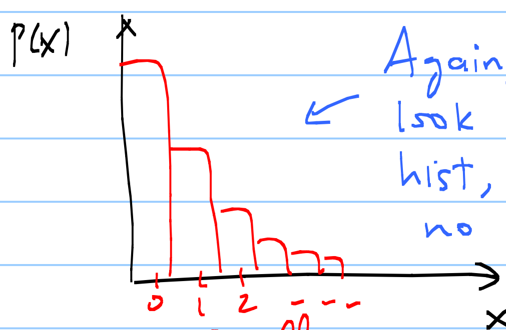
prob. of H on a single toss.

Note it is a mass function; $p(x) \geq 0$, $\sum_{x=0}^{n} p(x) = 1$

" it has parameters : n, π. [n = integers , $0 < \pi < 1$]

Look above for The meaning of The params.

Depending on The value of The params, it can look like



Small n

Again, note: look like hist, but no data.

large n

In Lab you'll see how These look for different π values.

4) <u>Normal/Gaussian</u> , x = cont.

E.g. weight, height, temperature, ...

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

π — 3.1415

Note: if $\mu=0, \sigma=1$ Then $f(x)$ is std. Normal.

parameters/meaning : μ, σ.

μ — measure of location or middle, or centrality.

σ — measure of spread

Irrelevant

Important: Resist the temptation to call μ and σ mean and standard deviation, atleast for now. otherwise you'll get very confused. They are simply parameters of the distribution.

e.g. between a, b

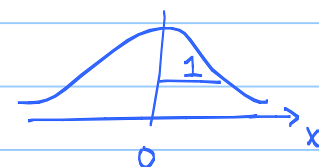
Recall: given a dist. we can find prop/prob That x is somewhere:

(Also note That props/probs are important and useful because They can be observed/measured in samples.)

→ For Normal (μ, σ):

$$= \int_a^b \underbrace{\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}_{= f(x) \text{ of Normal}} dx$$

→ For std. Normal: $\frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{1}{2}x^2} dx$



Note: std. Normal = Normal($\mu=0$, $\sigma=1$) = symmetric.

Unfortunately, integrals of This type can be done only numerically.

Next time!

Q1: Suppose x follows The exponential dist. (written $x \sim \exp(\lambda)$)

What is The prob That x will be exactly 2.67? Hint: prob = area.

1) $f(x=2.67)$ 2) $f(x=0)$, $\boxed{3)\ 0}$

$$\text{prob}(x=2.67) = \int_{2.67}^{2.67} \lambda e^{-\lambda x} dx =$$

Suppose $x \sim \text{Binom}(n, \pi)$, ie. $p(x) = \binom{n}{x} \pi^x (1-\pi)^{n-x}$.

What is The prob. That x will be exactly 1?

$\boxed{1)\ p(x=1)}$ 2) $p(x=0)$ 3) 0

For discrete/Categ. $\text{prob}(x=1) = p(x=1) = \binom{n}{1} \pi^1 (1-\pi)^{n-1} \neq 0$

hw_lect 4-1

The Bernoulli dist. discussed in The lecture does have a formula:

$p(x) = \pi^x (1-\pi)^{1-x}$, where $0<\pi<1$ is some param. and $x=0,1$.

a) Show that it's a mass function.

b) What's The prob. of getting $x=1$ (i.e. proportion of times do we expect to get $x=1$)?

---

hw_lect 4-2 Show that

a) $\int_0^\infty \lambda e^{-\lambda x}\,dx = 1$

b) $\sum_{x=0}^\infty \frac{e^{-\lambda}\lambda^x}{x!} = 1$ [Hint: use the Taylor series expansion for $e^{+\lambda}$]

c) $\int_{-\infty}^\infty \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = 1$ [use $\int_{-\infty}^\infty e^{-\frac{1}{2}x^2} dx = \sqrt{2\pi}$]
