Lecture 5 (Ch. 1)

Last time: Distribution, $f(x)$ or $p(x)$ = mathematical functions
(cont. → $f(x)$; Disc./Categ. → $p(x)$)

Dists ≠ hists

histogram of data = frequencies in data

But here are some "connections": (But, still, keep them separate!)

1) The population is described by a distribution.
2) The sample " " " " histogram.

If data come from a given population,
Then the hist of data will resemble the dist.

Some dists are named (e.g. Bernoulli, Binomial). They are famous because either they have desirable mathematical properties, or because there are lots of data in Nature whose hists look like these dists. Here are some:

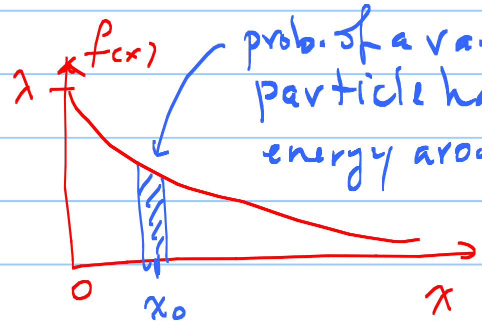
1) Exponential (family), x = continuous $\boxed{Exp(\lambda)}$

E.g. Energy of particles (Radioactivity), inter-arrival time, ...

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

Note the parameter: $\lambda > 0$

Meaning: $\lambda = \dfrac{1}{\text{"average } x\text{"}}$ (later).



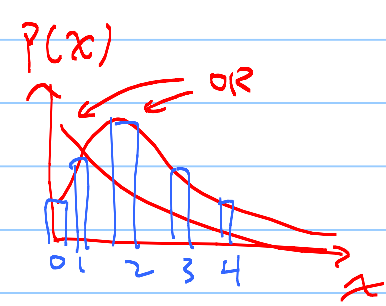
2) Poisson, x = discrete. $\boxed{Poiss(\lambda)}$

e.g. number of visits to a website, per hour.

$$\underbrace{p(x)}_{\text{prob. of } x \text{ visits per hour.}} = \frac{e^{-\lambda}\lambda^{x}}{x!}, \quad x = 0, 1, 2, \ldots$$

parameter: $\lambda > 0$

meaning: λ = "average x" (later)



3) Binomial (revisited) x = discrete $\boxed{\text{Binom}(n, \pi)}$

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.

# of Hs out of n tosses

prob. of H on a single toss.

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

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

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 the shape depends on the π param.

4) Normal/Gaussian , x = cont. $\boxed{N(\mu, \sigma)}$

E.g. temperature, $\overbrace{\text{height, weight}}^{\text{of one gender}}$, blood pressure, ...

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

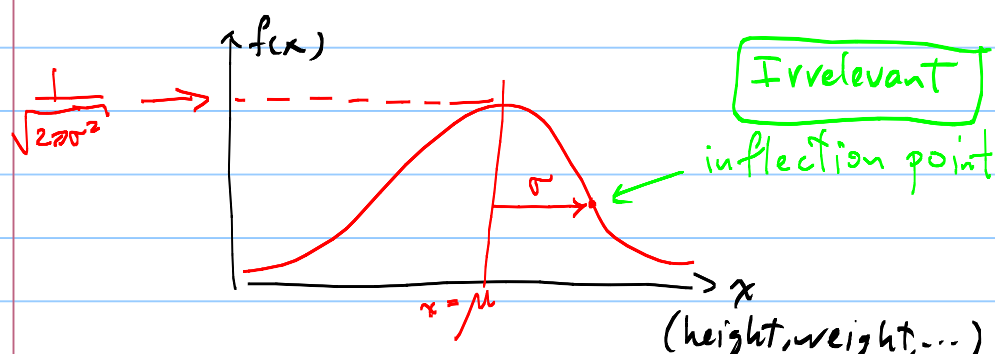
$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

π = 3.1415

parameters/meaning : μ, σ

μ: measure of location or middle, or centrality.

σ: measure of spread



Important: Resist the temptation to call μ, σ mean and standard deviation, atleast for a while; otherwise you'll get very confused. μ, σ are simply the params of the Normal dist. $N(\mu, \sigma)$ [some books write $N(\mu, \sigma^2)$]

So far, we have

Bernoulli
Binomial
Poisson

(later) Related.

} $P(x)$ = (probability) mass function (pmf)

$P(x)$: discrete/categ.

Uniform (in hw)
Exponential
Normal

} $f(x)$ = (probability) density function (pdf)

$f(x)$: continuous.

Recall that there is a connection between dists and probs.

⇒ For discrete/Categ. variables : prob $= \sum_{x} p(x)$

For example, if $x \sim \text{Binom}(n, \pi)$, then

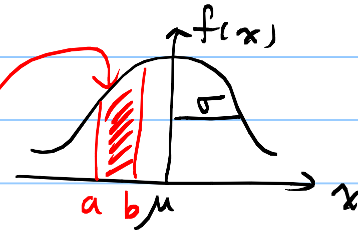
$$\text{prob}(a \le x \le b) = \sum_{x=a}^{b} \underbrace{\boxed{\binom{n}{x} \pi^{x} (1-\pi)^{n-x}}}_{p(x)}$$

Use either this formula or Table II.

⇒ For continuous variables : prob $= \int f(x)\,dx$ = area

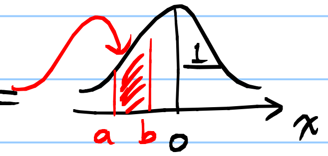
For example, if $X \sim N(\mu, \sigma)$, then

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



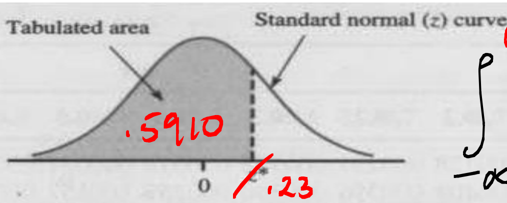
Let's start with the simpler std. Norm., i.e. $N(\mu=0, \sigma=1)$:

$$\text{pr}(a < x < b) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{1}{2}x^2}\,dx =$$



symmetric about zero

Unfortunately, such integrals have no closed-form. Approx. numerically.



$$\int_{-\infty}^{0.23} f(x)\,dx = 0.5910$$

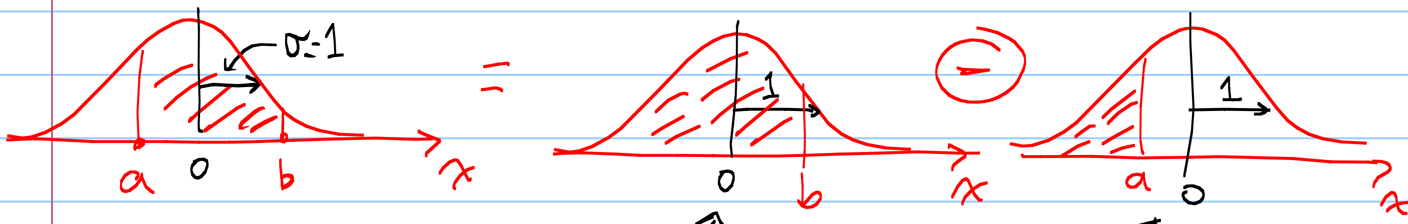
Table I gives "left area"

z*	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852

In 390, use Table I, unless the problem says "By R."

To find the area <u>between</u> 2 x's, there is a trick:

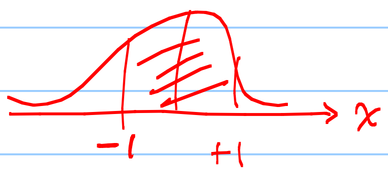
$\text{prob}(a < x < b) =$ area between a & b =



$= \text{prob}(x < b) - \text{prob}(x < a)$

Both of these can be obtained from Table I.

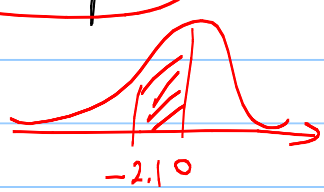
Example: What's the prob of $-1 < x < 1$ for $N(0,1)$?



$= .8413 - .1587 = 0.6826$

(famous 68% from High School)

Example: How about between -2.1 and 0?



$= 0.5 - .0179 = 0.4821$

Example: $\text{prob}(x > -2.1)$

$$\left.\begin{array}{l} 1-(0.0179) \\ \text{or } 0.4821+0.5 \end{array}\right\} = 0.9821$$

hw-lect 5-1 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 $\int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2}\,dx = \sqrt{2\pi}$]

---

hw-lect 5-2

Let x have a Normal dist. with params μ, σ, ie. $x \sim N(\mu, \sigma)$

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

a) Find The density function $(f(z))$ for $z = \frac{x-\mu}{\sigma}$.

$\boxed{\text{Hint}}$ Start with $\int_{-\infty}^{\infty} f(x)\,dx = 1$, do $\boxed{\text{not}}$ do The integral,

important

but instead, do a change of variable until you get $\int_{-\infty}^{\infty} [\,\cdots] \,dz = 1$. Then $[\ldots]$ is $f(z)$.

b) In $f(z)$, in The place where you would expect to find μ and σ, what numbers do you see?

---

hw-lect 5-3 Find The prob. of $X \le 2$ if

a) $X \sim$ Bernoulli (π= arbitrary)

b) $X \sim \exp(\lambda = 3)$

c) $X \sim$ poiss ($\lambda = 3$)

d) $X \sim$ Binom ($n = 10, \pi = \frac{1}{4}$)

e) $X \sim N(0,1)$