

Lecture 19

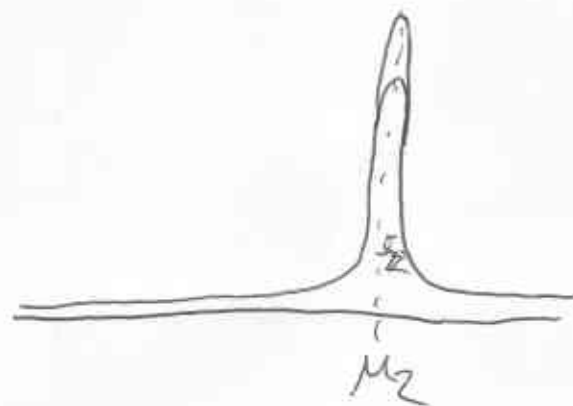
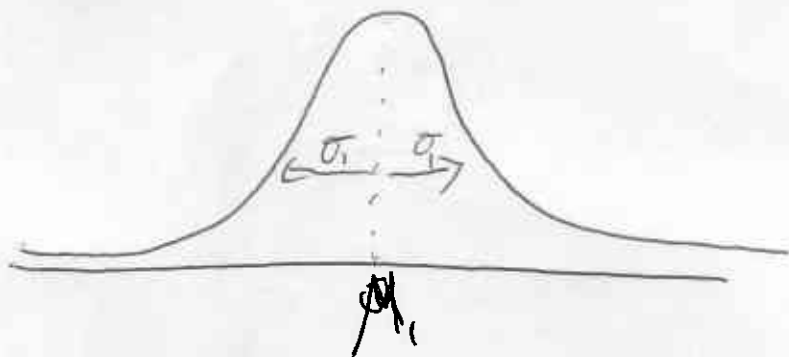
①

Normal Distribution

A continuous r.v. X has normal distribution with mean μ , standard deviation σ where $-\infty < \mu < \infty$ and $0 < \sigma$ if its pdf is

$$f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \left(\begin{array}{l} \text{sometimes written} \\ \text{as } \phi(x) \end{array} \right)$$

graphs of the pdf



μ controls location
 σ controls spread.

"Symmetric bell shaped"

Note μ is also the median for normal distribution

Notation

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

of course if $X \sim N(\mu, \sigma)$

$E[X] = \mu$

$Var(X) = \sigma^2$

$sd(X) = \sigma$

Applications of the normal

- many physical measurements, heights, weights, ...
- errors in measurements in scientific experiments
- The CLT (to be discussed later) gives that $\bar{X} \sim N(\mu, \sigma)$ as $n \rightarrow \infty$
- Normal distribution perhaps most useful in all of statistics

Special Case

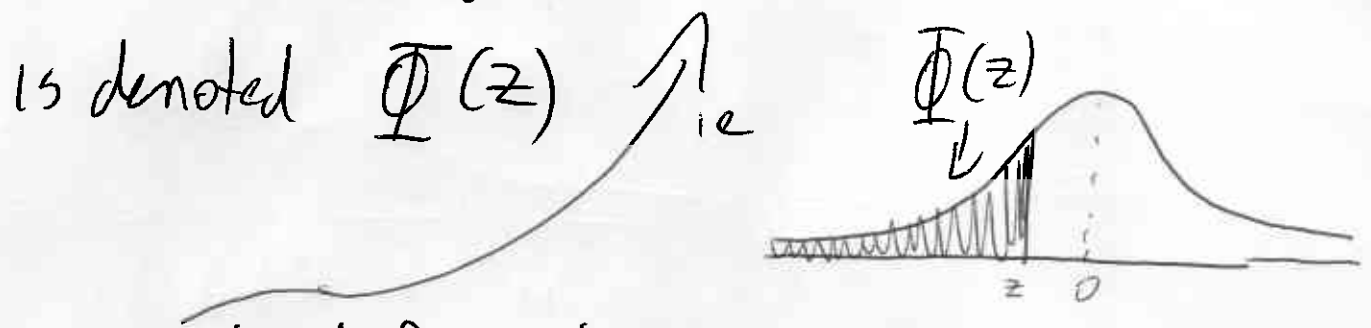
If $\mu=0$ and $\sigma=1$ then the distribution is called the standard normal distribution.

We usually represent a std normal r.v. using Z , so

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \quad -\infty < z < \infty$$

Note that the CDF of Z is

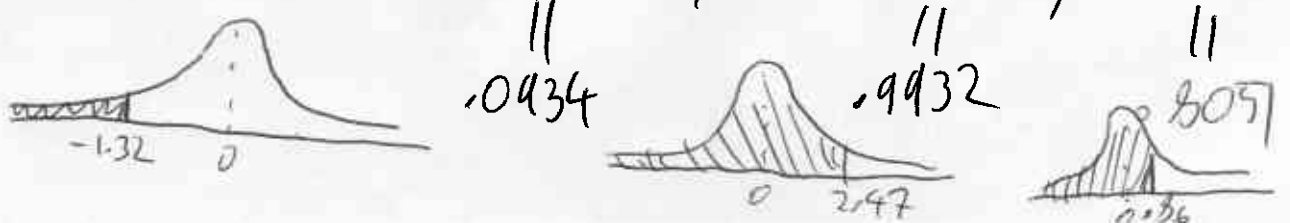
$$P(Z \leq z) = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy$$



no closed formula for this integral

instead we can use the table (see handout)

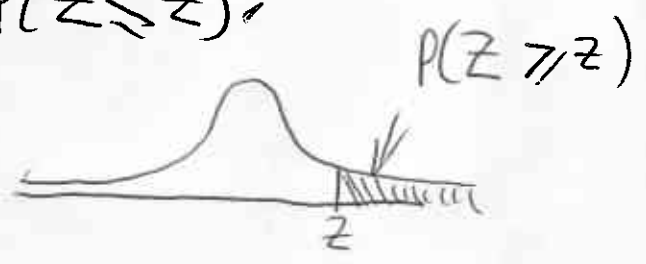
eg look up $\Phi(-1.32)$, $\Phi(2.47)$, $\Phi(0.87)$



Since

$\Phi(z) = P(Z \leq z)$ it is easy to look up directly probabilities of the form $P(Z \leq z)$.

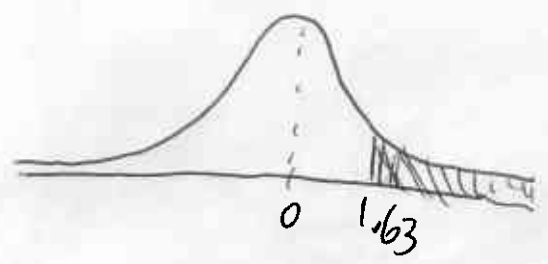
$P(Z > z) = ?$ 12



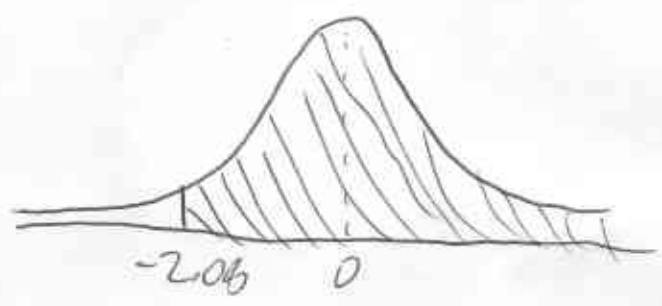
Recall from last time

$$P(Z > z) = 1 - F(z) = 1 - \Phi(z)$$

What is $P(Z > 1.63)$, $P(Z > -2.08)$

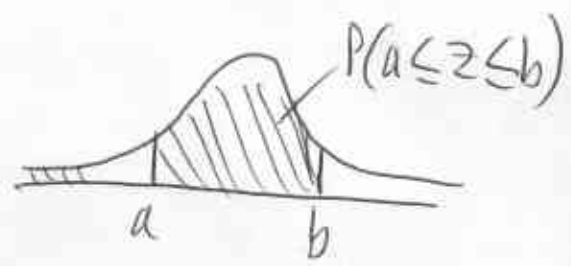


$$= 1 - .9484 = .0516$$



$$= 1 - .0168 = .9832$$

What about $P(a \leq Z \leq b)$?
 $a < b$

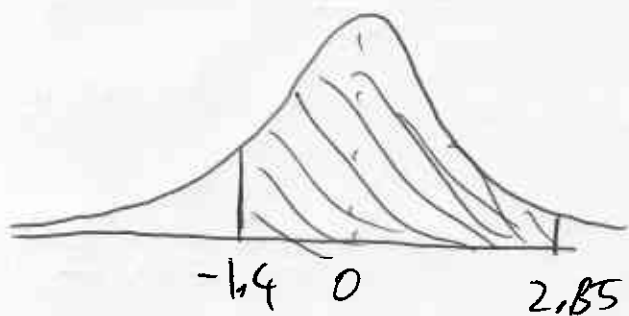


Recall from last time

$$P(a \leq Z \leq b) = F(b) - F(a) \\ = \Phi(b) - \Phi(a)$$

what is $P(-1.4 \leq z \leq 2.85)$

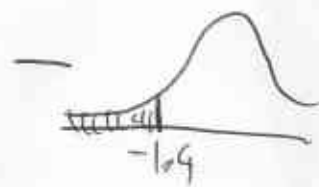
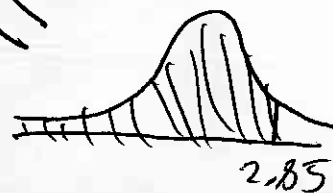
④



$$= \Phi(2.85) - \Phi(-1.4)$$

$$= .9978 - .0808$$

$$= .9170$$

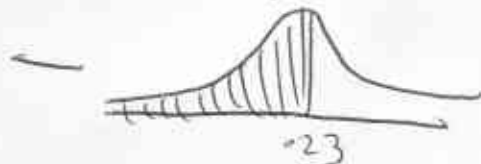
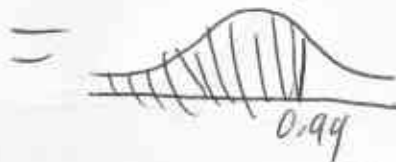
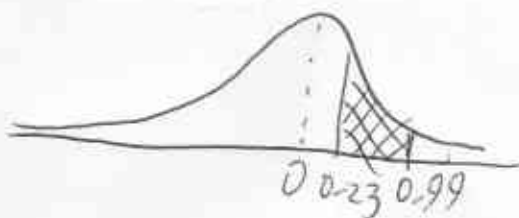


$P(0.23 \leq z \leq 0.99)$

$$= \Phi(0.99) - \Phi(0.23)$$

$$= .8389 - .5910$$

$$= .2479$$



what about $P(z \leq a \text{ or } z \geq b)$?

$$= P(z \leq a) + P(z \geq b)$$

$$= \Phi(a) + 1 - \Phi(b)$$



what if $\mu \neq 0$ and $\sigma \neq 1$?

5

If X is a r.v with normal distribution mean μ
sd σ then

$Z = \frac{X - \mu}{\sigma}$ is a r.v with std normal distribution

ie we transform our problem so that we deal
with a std normal r.v. and then look up
probability in table

eg suppose $X \sim N(100, 15)$
what is $P(80 \leq X \leq 105)$?

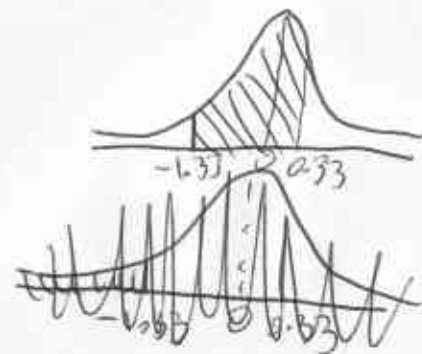
$$P\left(\frac{80 - 100}{15} \leq \frac{X - 100}{15} \leq \frac{105 - 100}{15}\right)$$

$$= P(-1.33 \leq Z \leq 0.33)$$

$$= \Phi(0.33) - \Phi(-1.33)$$

$$= .6293 - .0918$$

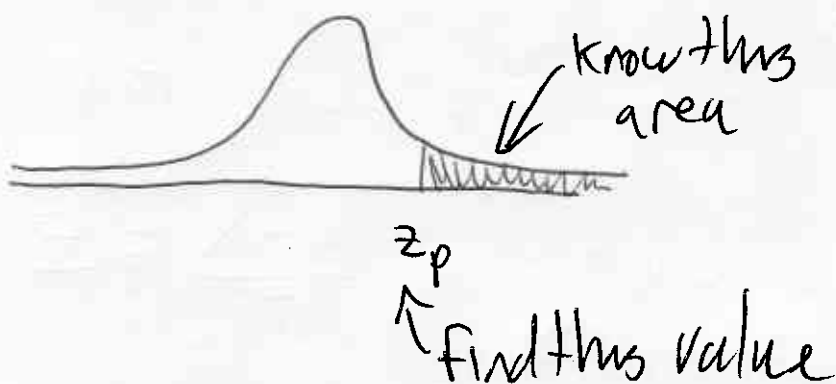
$$= .5375$$



How about the inverse problem?

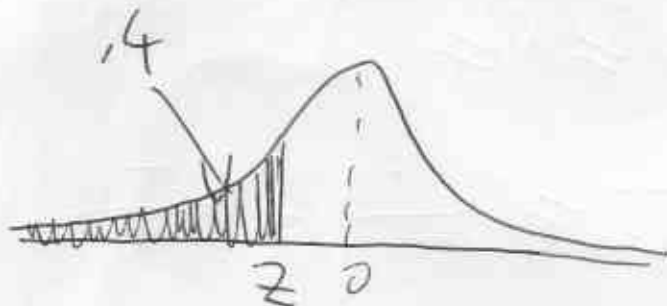
(6)

ie



Easy. Just use the table in the reverse order
ie look for the closest probability then
find z

eg what's z if $P(Z \leq z) = 0.4$



closest values are $\Phi(-0.25) = .4013$

and $\Phi(-0.26) = .3974$

so either use $z = -0.25$ or for
more accuracy linearly interpolate.

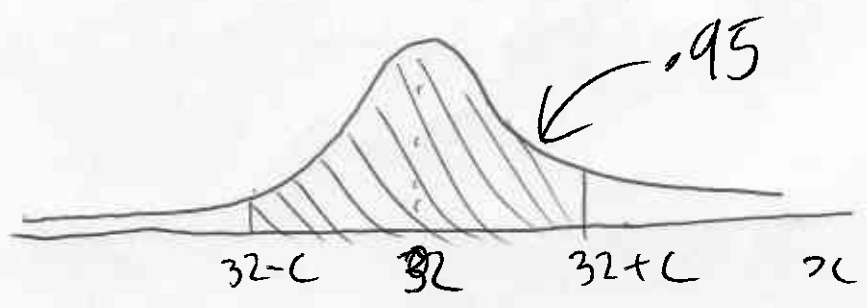
with non standard normal need to use

$$X = \sigma Z + \mu \text{ to unstandardize.}$$

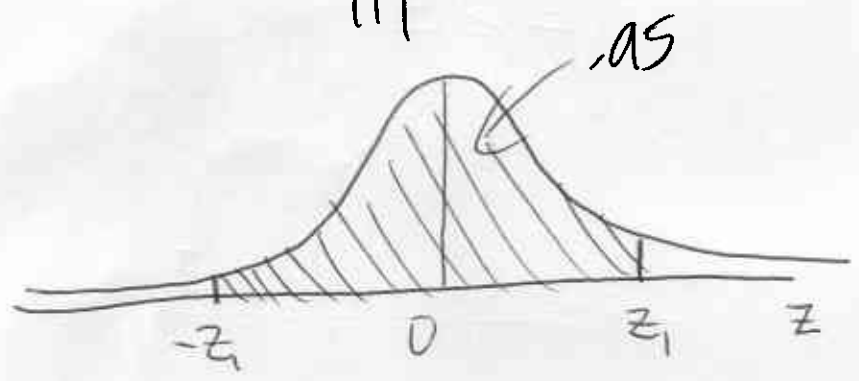
eg Suppose $X \sim N(32, 4)$

what value of c is

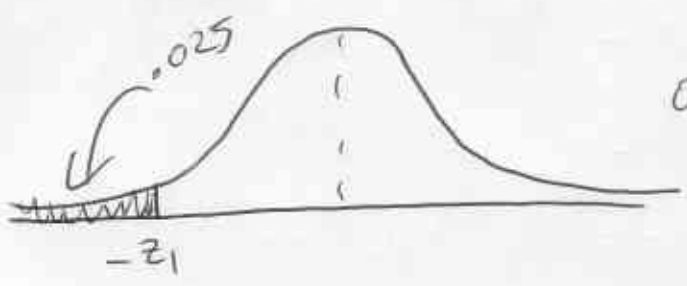
$$P(32 - c \leq X \leq c + 32) = .95$$



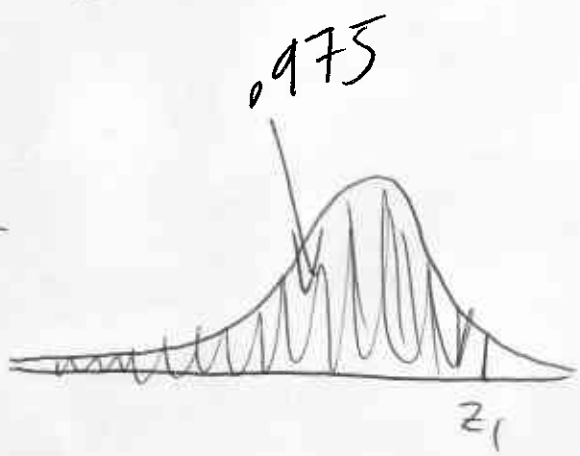
|||



so either



or



from the table

(9)

$$\Phi(-1.96) = .0250$$

$$\Phi(1.96) = .9750$$

$$x_1 = 4(-1.96) + 32 = 24.16$$

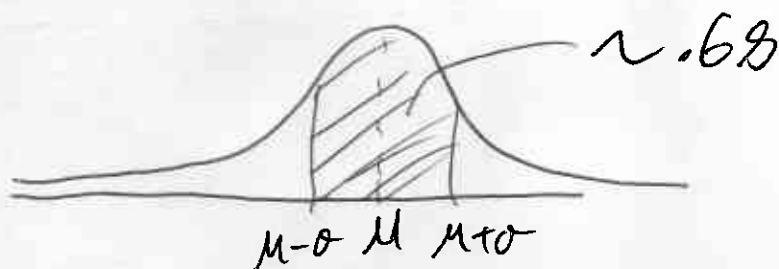
$$x_2 = 4(1.96) + 32 = 39.84$$

$$\text{i.e. } P(24.16 \leq X \leq 39.84) = .95$$

$$\text{note } c = 4(1.96) = 7.84$$

Some empirical rules about normal dist

1. Approximately 68% is within 1SD of the mean

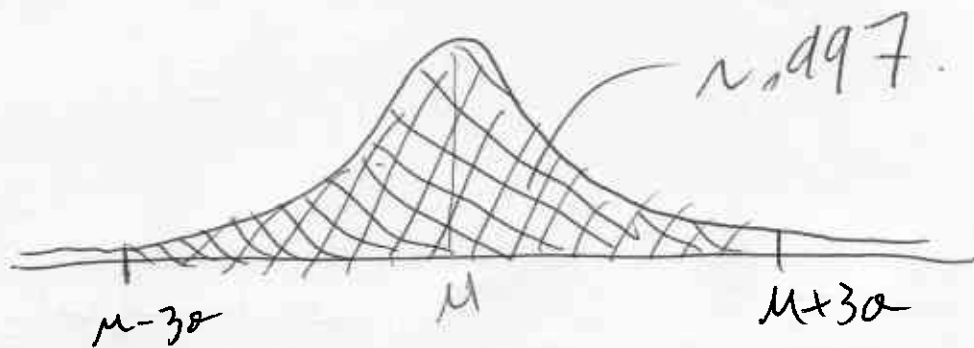


2. Approximately 95% is within 2SD of the mean



3. Approximately 99.7% is within 3SD of the mean

(9)



Normal approximation to the binomial

If X is a binomial rv n, p and not too skewed then X has approximately normal distribution $\mu = np$ $\sigma = \sqrt{np(1-p)}$

In practice approximation reasonable if $np \geq 10$ and $n(1-p) \geq 10$.