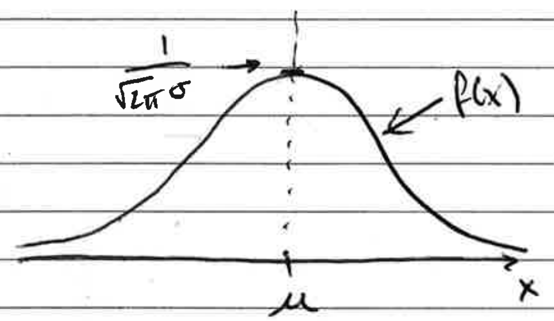
5.4 Normal Random Variable

A random variable X is a normal random variable (or X is normally-distributed) with parameters μ and σ^2 if its probability density function is

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

- recall, then $P\{a \le X \le b\} = \int_a^b f(x)\,dx$

- 

- "bell-shaped" curve

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

– see p. 169 in notes

we'll check these in the context of standard normal then generalize...

* Let's first check some properties of the standard normal random variable X with

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \qquad (\text{i.e. } \sigma = 1, \mu = 0)$$

$$-\infty < x < \infty$$

- First, let's confirm that

$$\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

is equal to one (is it?)

Let $I = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$

nifty trick...

Note:

$$I^2 = \left(\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \right) \left(\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy \right)$$

$$= \frac{1}{2\pi} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-(x^2+y^2)/2} dx\, dy$$

convert to polar coordinates $dx\,dy \rightarrow r\,dr\,d\theta$

$r^2 = x^2 + y^2$

$$I^2 = \frac{1}{2\pi} \int_0^{2\pi} \left[\int_0^{\infty} e^{-r^2/2} r\,dr \right] d\theta$$

$$= \frac{1}{2\pi} 2\pi \int_0^{\infty} r e^{-r^2/2} dr$$

$$= -e^{-r^2/2} \Big|_0^{\infty} = 0 - (-1) = 1 \checkmark$$

So $I^2 = 1 \Rightarrow I = 1$ (noting that $I > 0$)

- Next, let's compute the expected value $E[X]$.

$$E[X] = \int_{-\infty}^{\infty} x\, f(x)\, dx = \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x e^{-\frac{x^2}{2}} dx$$

$$= \frac{1}{\sqrt{2\pi}} \left[\int_{-\infty}^{0} x e^{-\frac{x^2}{2}} dx + \int_{0}^{\infty} x e^{-x^2/2} dx \right]$$

$$\int_{0}^{\infty} x e^{-\frac{x^2}{2}} dx = \int_{0}^{-\infty} -e^{+u} du = -e^{+u}\Big|_{0}^{-\infty} = +1$$

$u = -\frac{x^2}{2}$

$du = -x\,dx$

$$\int_{-\infty}^{0} x e^{-\frac{x^2}{2}} dx = \int_{-\infty}^{0} -e^{u} du = -e^{u}\Big|_{-\infty}^{0} = -1$$

So both of these integrals are finite and it follows

$$\boxed{E[X] = 0}$$

- Next, let's compute the Variance (X = standard normal)

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

$$E[X^2] = \int_{-\infty}^{+\infty} \frac{x^2}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^2 e^{-\frac{x^2}{2}} dx$$

$u = x \qquad du = dx$

$dv = x e^{-x^2/2} dx \qquad v = -e^{-x^2/2}$

$$= \frac{1}{\sqrt{2\pi}} \left[\cancel{-x e^{-x^2/2} \Big|_{-\infty}^{\infty}}^{\,0} - \int_{-\infty}^{\infty} -e^{-x^2/2} dx \right]$$

$$= \underbrace{\boxed{\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2} dx}}_{=1} = \int_{-\infty}^{\infty} f(x) dx = 1$$

So $Var(X) = 1 - 0 = 1$

$\therefore$ For standard normal distribution $P\{X \in B\} = \int_B f(x) dx$

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

$E[X] = 0$

$Var(X) = 1$

Cumulative Distribution Function

/ common tabulated quantity – see p. 190 (for $x>0$)

$$\boxed{F(x) = P\{\Xi \le x\} = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}\,dt \quad \equiv \Phi(x) \text{ often}}$$

← Ξ = standard normal distribution

Note:

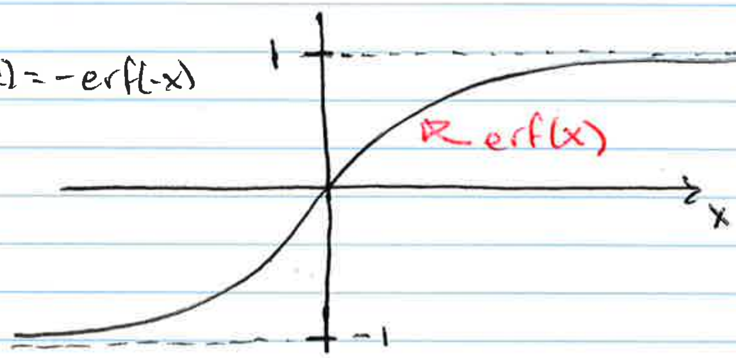
$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-s^2}\,ds = \text{error function}$$

$$\mathrm{erfc}(x) = \frac{2}{\sqrt{\pi}}\int_x^{\infty} e^{-s^2}\,ds = \text{complementary error function}$$

$\mathrm{erf}(0) = 0$, $\mathrm{erf}(x) = -\mathrm{erf}(-x)$

$\mathrm{erf}(x\to\infty) = 1$

$\mathrm{erf}(x) + \mathrm{erfc}(x) = 1$



Let $t = -\sqrt{2}\,s$

$dt = -\sqrt{2}\,ds$

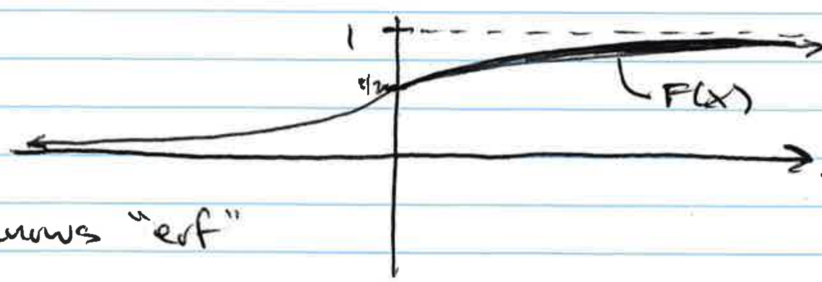
then

$$F(x) = \int_{+\infty}^{-\frac{x}{\sqrt{2}}} -\frac{1}{\sqrt{2\pi}} e^{-s^2}\sqrt{2}\,ds = \frac{1}{\sqrt{\pi}}\int_{+\infty}^{-\frac{x}{\sqrt{2}}} -e^{-s^2}\,ds = \frac{1}{\sqrt{\pi}}\int_{-\frac{x}{\sqrt{2}}}^{\infty} e^{-s^2}\,ds$$

So

$$F(x) = \frac{1}{2}\,\mathrm{erfc}\left(-\frac{x}{\sqrt{2}}\right)$$

$$= \frac{1}{2}\left(1 - \mathrm{erf}\left(-\frac{x}{\sqrt{2}}\right)\right) = \overbrace{\frac{1}{2}\left(1 + \mathrm{erf}\left(\frac{x}{\sqrt{2}}\right)\right)}^{\Phi(x)}$$



The table in back p. 190 shows values for this

Matlab knows "erf"

*

We can connect normally distributed X with standard normal random variable Y as follows.

- Let X be a normal random variable with parameters μ and σ^2. So

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

- Let $\boxed{Y = \frac{X-\mu}{\sigma}}$. Then

often called $Z = \frac{X-\mu}{\sigma}$

$$F_Y(y) = P\{Y \le y\} = P\left\{\frac{X-\mu}{\sigma} \le y\right\} = P\{X \le y\sigma + \mu\}$$

$$= \int_{-\infty}^{y\sigma+\mu} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,dx$$

but then

$$f_Y = \frac{d}{dy} F_Y(y) = \frac{d}{dy} \int_{-\infty}^{y\sigma+\mu} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,dx$$

$$= \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \Bigg|_{x = y\sigma+\mu} \cdot \underbrace{\sigma}_{\frac{d}{dy}(y\sigma+\mu)}$$

$$= \frac{1}{\sqrt{2\pi}} e^{-y^2/2}$$

← standard normal prob. density function.

So if $E[Y] = 0$

$$Var(Y) = E[Y^2] - (E[Y])^2 = 1 - 0 = 1$$

then

$$E[X] = E[\mu + \sigma Y]$$

$$= \int_{-\infty}^{+\infty} (\mu + \sigma y) f_Y(y)\, dy$$

$$= \mu \underbrace{\int_{-\infty}^{+\infty} f_Y(y)\, dy}_{=1} + \sigma \underbrace{\int_{-\infty}^{+\infty} y f_Y(y)\, dy}_{=0 \; = E[Y]} = \mu \checkmark$$

and

$$E[X^2] = E[(\mu + \sigma Y)^2]$$

$$= E[\mu^2 + 2\mu\sigma Y + \sigma^2 Y^2]$$

$$= \int_{-\infty}^{+\infty} (\mu^2 + 2\mu\sigma y + \sigma^2 y^2) f_Y(y)\, dy$$

$$= \mu^2 \underbrace{\boxed{\int_{-\infty}^{+\infty} f_Y(y)\, dy}}_{=1} + 2\mu\sigma \underbrace{\boxed{\int_{-\infty}^{+\infty} y f_Y(y)\, dy}}_{= E[Y] = 0} + \sigma^2 \underbrace{\boxed{\int_{-\infty}^{+\infty} y^2 f_Y(y)\, dy}}_{E[Y^2] = 1}$$

$$= \mu^2 + \sigma^2$$

so

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

$$= \mu^2 + \sigma^2 - \mu^2$$

$$= \sigma^2 \checkmark$$

EX

Suppose X is a normal random variable with $\mu = -3$ and $\sigma^2 = 9$

Find $P\{-4 \le X \le 0\}$

write $Y = \dfrac{X-\mu}{\sigma}$

then $X = \mu + \sigma Y$

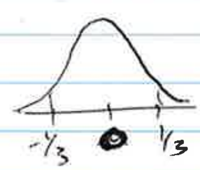
$$P\{-4 \le X \le 0\} = P\{-4 \le \mu + \sigma Y \le 0\}$$

$$= P\left\{\frac{-4-\mu}{\sigma} \le Y \le \frac{0-\mu}{\sigma}\right\}$$

$$= P\left\{\frac{-4-(-3)}{\sqrt{9}} \le Y \le \frac{0-(-3)}{\sqrt{9}}\right\}$$

$$= P\left\{-\tfrac{1}{3} \le Y \le 1\right\}$$

$$= P\{Y \le 1\} - \underbrace{P\{Y \le -\tfrac{1}{3}\}}_{\text{by symmetry}}$$



$$= P\{Y \ge \tfrac{1}{3}\}$$

$$= P\{Y \le 1\} - P\{Y \ge 1/3\}$$

$$= P\{Y \le 1\} - \left(1 - P\{Y < 1/3\}\right)$$

see p. 167

$\Phi(1) - 1 + \Phi(1/3)$

$$= F(1) - 1 + F(1/3) \qquad \text{where } F(x) = \frac{1+\operatorname{erf}\left(\frac{x}{\sqrt{2}}\right)}{2}$$

$$= \frac{1+\operatorname{erf}\left(\frac{1}{\sqrt{2}}\right)}{2} - 1 + \frac{1+\operatorname{erf}\left(\frac{1}{3\sqrt{2}}\right)}{2} = 0.8413 - 1 + 0.6306 = \boxed{0.4719}$$

*

Note:

$$P\left\{-\tfrac{1}{3} \le I \le 1\right\}$$

$$= P\{I \le 1\} - P\{I \le -1/3\}$$

$$= \frac{1}{2}\left(1 + \operatorname{erf}\left(\frac{1}{\sqrt{2}}\right)\right) - \frac{1}{2}\left(1 + \operatorname{erf}\left(\frac{-1/3}{\sqrt{2}}\right)\right)$$

$$= \frac{1}{2} + \frac{1}{2}\operatorname{erf}\left(\frac{1}{\sqrt{2}}\right) - \frac{1}{2} - \frac{1}{2}\operatorname{erf}\left(\frac{-1/3}{\sqrt{2}}\right)$$

$$= \frac{1}{2}\left[\operatorname{erf}\left(\frac{1}{\sqrt{2}}\right) - \operatorname{erf}\left(\frac{-1/3}{\sqrt{2}}\right)\right]$$

Note $\operatorname{erf}\left(\frac{-1/3}{\sqrt{2}}\right) = -\operatorname{erf}\left(\frac{1/3}{\sqrt{2}}\right)$

$$= \frac{1}{2}\left[\operatorname{erf}\left(\frac{1}{\sqrt{2}}\right) + \operatorname{erf}\left(\frac{1/3}{\sqrt{2}}\right)\right]$$

$$= \frac{1}{2}\left[0.6827 + 0.2611\right] = \underline{\underline{0.4719}}$$

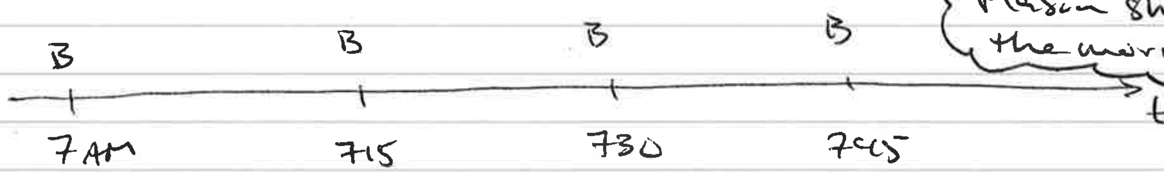
EXAMPLE - uniform distribution (Ex 3c, p. 185)

Buses arrive at a bus stop at 15 minute intervals starting at 7AM. (Arrive at 7AM, 7:15AM, 7:30AM, ...)

If a passenger arrives at the stop at a time that is uniformly distributed between 7 and 7:30 what is the probability that the passenger waits ...

a) less than 5 minutes for a bus?

Think of this in terms of metro to Mason shuttles in the morning



passenger arrives in this interval with uniform distribution,

Let T = uniform random variable on $[7, 7:30]$

↳ convert to minutes past 7 $[0, 30]$

$$f(x) = \begin{cases} \dfrac{1}{30} & 0 \le x \le 30 \\ 0 & \text{otherwise} \end{cases}$$

~~less than 5 min wait~~

$\cancel{P\{7:10 < T < 7:25\} + P\{7:25 < T < 7:30\}} =$

... less than 5 min wait ...

$$P\{10 < T < 15\} + P\{25 < T < 30\}$$

$$= \int_{10}^{15} \frac{1}{30}\,dx + \int_{25}^{30} \frac{1}{30}\,dx$$

$$= \frac{15-10}{30} + \frac{30-25}{30} = \frac{5}{30} + \frac{5}{30} = \frac{10}{30} = \boxed{1/3}$$

b) more than 10 minutes for a bus?

$$P\{0 < T < 5\} + P\{15 < T < 20\}$$

$$= \int_{0}^{5} \frac{1}{30}\,dx + \int_{15}^{20} \frac{1}{30}\,dx$$

$$= \frac{5}{30} + \frac{5}{30} = \frac{10}{30} = \boxed{1/3}$$

Metro to Mason has this same schedule ... shuttles run every 15 minutes (at least in the morning...)

Problem — Normal Distribution

$\boxed{5.21}$

Suppose that the height (in inches) of a 25-year-old man is a normal random variable with parameters $\mu = 71$ and $\sigma^2 = 6.25$ (note $\sigma = 2.5 \Rightarrow \sigma^2 = 6.25$)

a) What percentage of 25-year-old men are over 6 ft, 2 inches tall? $= 6 \cdot 12 + 2 = 74$ inches

X = height (normal random variable)

let $Z = \dfrac{X - \mu}{\sigma} = \dfrac{X - 71}{2.5}$

so $P\{X \geq 74\} = P\left\{\dfrac{X-71}{2.5} \geq \dfrac{74-71}{2.5}\right\}$

$= P\left\{Z \geq \dfrac{3}{2.5}\right\} = 1 - P\left\{Z \leq \dfrac{6}{5}\right\}$

$\dfrac{3}{\left(\frac{5}{2}\right)}$

$= 1 - 0.8849$ ← From Table, p.190

$= 0.1151$

or $= 1 - \frac{1}{2}\left(1 + \operatorname{erf}\left(\dfrac{6/5}{\sqrt{2}}\right)\right) = 0.1151$

using $P\{Z < z\} = \frac{1}{2}\left(1 + \operatorname{erf}\left(\dfrac{z}{\sqrt{2}}\right)\right)$ see notes p. 167

b) What percentage of 25-year old men who are ~~taller than~~ 6ft or taller, are taller than 6ft 5 inches?

This is a conditional probability

$$P\{X > 77 \mid X \geq 72\}$$

$$= \frac{P\{X > 77 \text{ AND } X \geq 72\}}{P\{X \geq 72\}} = \frac{P\{X > 77\}}{P\{X \geq 72\}}$$

$$= \frac{P\{Z > \frac{77-71}{2.5}\}}{P\{Z \geq \frac{72-71}{2.5}\}}$$

recall

$$\boxed{P(E|F) = \frac{P(EF)}{P(F)}}$$

conditional probability
p.57

$$= \frac{P\{Z > \frac{6}{2.5}\}}{P\{Z \geq \frac{1}{2.5}\}} = \frac{P\{Z > \frac{12}{5}\}}{P\{Z \geq \frac{2}{5}\}}$$

$$= \frac{1 - P\{Z \leq 2.4\}}{1 - P\{Z < 0.4\}} = \frac{1 - \frac{1}{2}\left(1 + \text{erf}\left(\frac{2.4}{\sqrt{2}}\right)\right)}{1 - \frac{1}{2}\left(1 + \text{erf}\left(\frac{0.4}{\sqrt{2}}\right)\right)}$$

$$= \frac{1 - 0.9918}{1 - 0.6554} \approx \frac{0.0082}{0.3446} \approx 0.0238$$

c) $P\{X > 78 \mid X \geq 72\} = \ldots = \frac{1 - P\{Z < 2.8\}}{1 - P\{Z < 0.4\}} = \frac{0.0026}{0.3446} \approx 0.0074$

5.4.1 Normal Approximation of Binomial Distribution

De Moivre - Laplace Limit Theorem

see Section 4.6 pp.127 →

Recall the Binomial (Discrete) Random Variable, X_B

→ probability mass function

$$p(i) = \binom{n}{i} p^i (1-p)^{n-i} \qquad i=0,1,\ldots,n$$

= probability of $X_B = i$ successes in n independent trials

(where p = probability of success, $1-p$ = prob. of failure)

→ Expected Value: $E[X_B] = np$

→ Variance : $Var(X_B) = np(1-p)$

→ Standard Deviation $SD = \sqrt{Var(X_B)} = \sqrt{np(1-p)}$

Now ~~[illegible]~~, we "standardize" X_B by writing

$$Z_B = \frac{X_B - np}{\sqrt{np(1-p)}}$$

~~[illegible]~~

$E[Z_B] = 0$

$Var(Z_B) = 1$

$$E[X_B] = E\left[np + \sqrt{} \cdot Z_B\right]$$

$$= np + \sqrt{}\, E[Z_B] \quad \Rightarrow E[Z_B] = 0$$

verify...

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

$$= E\left[(np)^2 + 2\sqrt{}\, np\, Z_B + (np(1-p)) Z_B^2\right] - (E[X_B])^2$$

$$np(1-p) = \cancel{(np)^2} - 0 + np(1-p) E[Z_B^2] - \cancel{(np)^2} \quad \Rightarrow E[Z_B^2] = 1$$

$\hookrightarrow Var(Z_B) = 1$

The DeMoivre-Laplace Limit Theorem says...

... the distribution function for Z_B

$$P\{Z_B \le z\} = P\left\{\frac{X_B - np}{\sqrt{np(1-p)}} \le z\right\}$$

for the discrete binomial random variable Z_B (standardized version)

converges to the continuous standard normal distribution function as $n \to \infty$. That is,

<u>DeMoivre-Laplace Limit Theorem</u>

Let X_B denote the number of successes that occur when n independent trials, each with probability of success p, are performed (i.e. discrete binomial R.V. X_B)

Then, for any $a < b$

$$P\left\{a \le \frac{X_B - np}{\sqrt{np(1-p)}} \le b\right\} \to \Phi(b) - \Phi(a)$$

as $n \to \infty$. Here

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt = P\{Z \le x\} = F(x) = \frac{1}{2}\left(1 + \operatorname{erf}\left(\frac{x}{\sqrt{2}}\right)\right)$$

where Z is a continuous standard normal R.V.

= Cumulative Distribution Function for Continuous Standard Normal Random Variable Z

EX

Let X = # of heads that occurs when a fair coin is flipped 64 times. Find $P\{X=32\}$

Binomial R.V.

$$P\{X=32\} = \binom{64}{32}\left(\tfrac{1}{2}\right)^{32}\left(1-\tfrac{1}{2}\right)^{32} = \binom{64}{32}\left(\tfrac{1}{2}\right)^{32}\left(\tfrac{1}{2}\right)^{32}$$

$$= \frac{64!}{32!\,32!}\,\frac{1}{2^{64}} \approx 0.099347$$

$\boxed{n=64,\ p=1/2}$

see BINOMDIST(k; n; p; mode) in Open Office ← (32; 64; 0.5; 0)

k: # of successes; n: # of trials; p: success probability (mode=0)

BINOMDIST(32, 64, 0.5, false) in EXCEL

Compare to **Normal Distribution** (continuous R.V.) $Z = \frac{X-32}{4}$

$np = 64 \cdot \frac{1}{2} = 32$

$np(1-p) = 64 \cdot \frac{1}{2} \cdot \frac{1}{2} = 16$

$\sqrt{np(1-p)} = 4$

$$P\{X=32\} \approx P\{31.5 \le X_N \le 32.5\}$$

$$= P\left\{\frac{31.5-32}{4} \le Z \le \frac{32.5-32}{4}\right\}$$

$$= P\left\{\frac{-0.5}{4} \le Z \le \frac{0.5}{4}\right\}$$

$$= P\left\{-\tfrac{1}{8} \le Z \le \tfrac{1}{8}\right\}$$

$$= \Phi\left(\tfrac{1}{8}\right) - \Phi\left(-\tfrac{1}{8}\right)$$

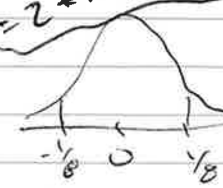
$$= \Phi\left(\tfrac{1}{8}\right) - \left[1 - \Phi\left(\tfrac{1}{8}\right)\right]$$

$$= 2\Phi\left(\tfrac{1}{8}\right) - 1 = 2\left(\tfrac{1}{2}\left(1+\operatorname{erf}\left(\tfrac{1/8}{\sqrt{2}}\right)\right)\right) - 1 = \operatorname{erf}\left(\tfrac{1/8}{\sqrt{2}}\right)$$

$= \boxed{0.099476}$

EXCEL: 2* NORMSDIST(0.125)-1

Open Office: = 2* NORMDIST(0.125;0;1;1)-1 = 2* NORMSDIST(0.125)-1

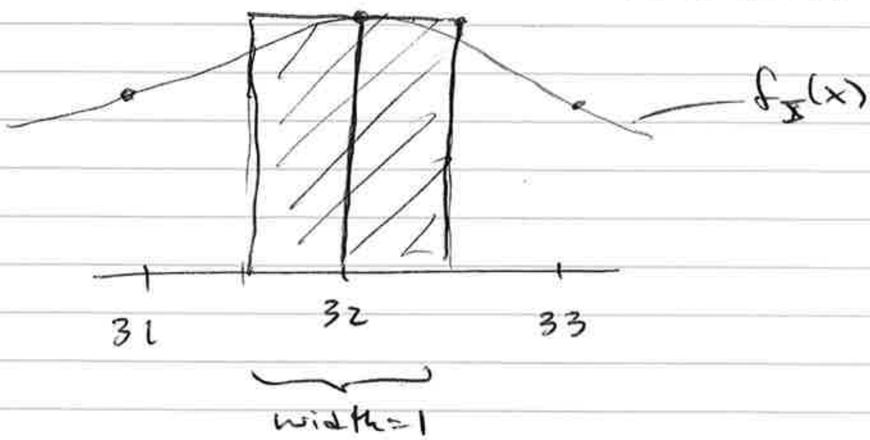


$\frac{1}{8} = 0.125$

Note the approximation $= \int_{32-1/2}^{32+1/2} f_X \, dx$

$$\underbrace{p(32) = P\{X = 32\}}_{\text{Binomial}} \approx \underbrace{P\{32 - 1/2 \le X \le 32 + 1/2\}}_{\text{Normal Distribution}} \approx f_X(32) * \Delta x$$

$\Delta x = 1$



height of probability mass function $* 1$

$\approx$ area under normal curve with width $= 1$

i.e. $p(32) \approx f_X(32) \cdot \underbrace{\Delta x}_{=1} = P\{32 - 1/2 \le X \le 32 + 1/2\}$

$\| \; P\{X = 32\}$

i.e. not $P\{X = 32\}$

need window of width $= 1$.

EX (Problem $\boxed{5.27}$)

In 10,000 tosses of a coin, the coin landed on heads 5800 times. Is it reasonable to assume the coin is not fair?

- Suppose the coin was fair and compute the probability of getting at least 5800 heads.
- Let X_B = # of heads

Binomial (for fair coin)
$n = 10,000$
$p = 1/2$

$E[X] = np = 5000$
$Var(X) = np(1-p) = \frac{10000}{2 \cdot 2}$
$= 2500$

$$P\{X_B \geq 5800\} \leftarrow$$

$$\approx P\{X_N \geq 5799.5\}$$

$$= P\left\{\frac{X_N - 5000}{\sqrt{2500}} \geq \frac{5799.5 - 5000}{\sqrt{2500}}\right\}$$

$$= P\left\{Z_N \geq \frac{799.5}{50}\right\}$$

$$= P\{Z_N \geq 15.99\}$$

$$= 1 - P\{Z_N < 15.99\} = 1 - \Phi(15.99)$$

$$= 1 - \frac{1}{2}\left(1 + \operatorname{erf}\left(\frac{15.99}{\sqrt{2}}\right)\right)$$

$$= 0 \quad \text{to within } 10^{-16}$$

So the 5800 heads with a fair coin seems highly unlikely.

EX (Problem $\boxed{5,28}$)

Suppose that 12% of the population is left-handed. Approximate the probability that there are at least 20 left-handed students in a school of 200 students.

State assumptions...

Let X = # of left handed students in school

Discrete Binomial R.V.

$$P\{X=20\} = \binom{200}{20}(0.12)^{20}(0.88)^{180}$$

lefty righty

← Probability of exactly 20 leftys.

~~approximate with normal dist.~~

Discrete Binomial

$$= \sum_{i=20}^{200}\binom{200}{i}(0.12)^i (0.88)^{200-i}$$

Probability of at least 20 lefty's ...

$$P\{X \geq 20\} \approx P\{X_N \geq 19.5\}$$

$$= P\left\{\frac{X_N - 200(0.12)}{\sqrt{200(.12)(.88)}} \geq \frac{19.5 - 200(0.12)}{\sqrt{200(.12)(.88)}}\right\}$$

$$= P\left\{Z_N \geq \frac{19.5 - 24}{4.596}\right\}$$

$$= P\{Z_N \geq -0.9791\}$$

$$= P\{Z_N \leq 0.9791\}$$

$$= \frac{1}{2}\left(1 + \operatorname{erf}\left(\frac{0.9791}{\sqrt{2}}\right)\right)$$

$$= 0.8362$$

So 83% chance there are at least 20 leftys in school with 200 students.