BITS Pilani
Pilani Campus

# Probability Distribution

Akanksha Bharadwaj
Asst. Professor, BITS Pilani

innovate achieve lead

BITS Pilani
Pilani Campus

# SS ZG536, ADV STAT TECHNIQUES FOR ANALYTICS
# Contact Session 4

# Binomial Random Variable

- A binary random variable
- e.g., head or tail in each toss of a coin; defective or not defective light bulb
- Generally called “success” and “failure”
- Probability of success is p, probability of failure is 1 – p

# Binominal Probability Distribution

An experiment for which Conditions 1–4 are satisfied is called a **binomial experiment.**

**1.** The experiment consists of a sequence of *n* smaller experiments called *trials,* where *n* is fixed in advance of the experiment.

**2.** Each trial can result in one of the same two possible outcomes, Head or Tails

**3.** The trials are **independent**, so that the outcome on any particular trial does not influence the outcome on any other trial.

**4.** The probability of success *P*(*S*) is constant from trial to trial; we denote this probability by *p.*

P(Heads) = 1/2

success

# Example

- The same coin is tossed successively and independently *n* times. We arbitrarily use *S* to denote the outcome *H* (heads) and *F* to denote the outcome *T* (tails)

$n = 3$

possible outcomes ?

$2^3 = 8$

HHH
HHT
HTH
HTT
THH
THT
TTH
TTT

if $n = 5$
then possible cases
$2^5$

# Exercise

- Suppose a certain city has 50 licensed restaurants, of which 15 currently have at least one serious health code violation and the other 35 have no serious violations. There are five inspectors, each of whom will inspect one restaurant during the coming week. The name of each restaurant is written on a different slip of paper, and after the slips are thoroughly mixed, each inspector in turn draws one of the slips *without replacement.*
- ***Is it a binomial experiment?***

HCV
no HCV
possible outcomes

NO trials are not independent

∵ w/o replacement

# Solution

$$P(S \text{ on first trial}) = \frac{35}{50} = .70$$

and

$$\begin{aligned} P(S \text{ on second trial}) &= P(SS) + P(FS) \\ &= P(\text{second } S \mid \text{first } S)P(\text{first } S) \\ &\quad + P(\text{second } S \mid \text{first } F)P(\text{first } F) \\ &= \frac{34}{49}\cdot\frac{35}{50} + \frac{35}{49}\cdot\frac{15}{50} = \frac{35}{50}\left(\frac{34}{49}+\frac{15}{49}\right) = \frac{35}{50} = .70 \end{aligned}$$

Similarly, it can be shown that $P(S \text{ on } i\text{th trial}) = .70$ for $i = 3, 4, 5$. However,

$$P(S \text{ on fifth trial} \mid SSSS) = \frac{31}{46} = .67$$

whereas

$$P(S \text{ on fifth trial} \mid FFFF) = \frac{35}{46} = .76$$

# Exercise

- A certain state has 500,000 licensed drivers, of whom 400,000 are insured. A sample of 10 drivers is chosen without replacement. The *i*th trial is labeled *S* if the *i*th driver chosen is insured.?
- Although this situation would seem identical to that of previous example, the important difference is that the size of the population being sampled is very large relative to the sample size.
- ***Is it a binomial experiment?***

Yes

insured

not insured

Let, 1st trial be success

$$P(S \text{ on } 2^{nd} \mid S \text{ on } 1^{st}) = \frac{3,99,999}{4,99,999} \approx 0.8$$

$$P(S \text{ on } 10^{th} \mid S \text{ on previous all trials}) = \frac{3,99,991}{4,99,991} = 0.7999$$

$$\approx 0.8$$

$$P(\text{Success}) \text{ or } P(\text{insured}) = 0.8$$

i.e. constant from trial to trial

- These calculations suggest that although the trials are not exactly independent, the conditional probabilities differ so slightly from one another that for practical purposes the trials can be regarded as independent with constant probability.
- Thus, to a very good approximation, the experiment is binomial with n=10 and p=.8

# Example

- Suppose, for example, that n=3 . Then there are eight possible outcomes for the experiment:

  SSS,SSF,SFS,SFF,FSS,FSF,FFS,FFF

  random variable (no. of success in each trial)

  $X = \{0, 1, 2, 3\}$

- From the definition of *X*,*X(SSF)*=2, X(SFF)=1, and so on. Possible values for *X* in an *n*-trial experiment are x=0,1,2,3….n .

- *Because the pmf of a binomial rv X depends on the two parameters n and p, we denote the pmf by b(x; n, p).*

$P(S) = p \qquad P(F) = (1-p)$

**Table 3.1** Outcomes and Probabilities for a Binomial Experiment with Four Trials

| Outcome | $x$ | Probability | Outcome | $x$ | Probability |
|---|---|---|---|---|---|
| $SSSS$ | 4 | $p^4$ | $FSSS$ | 3 | $p^3(1-p)$ |
| $SSSF$ | 3 | $p^3(1-p)$ | $FSSF$ | 2 | $p^2(1-p)^2$ |
| $SSFS$ | 3 | $p^3(1-p)$ | $FSFS$ | 2 | $p^2(1-p)^2$ |
| $SSFF$ | 2 | $p^2(1-p)^2$ | $FSFF$ | 1 | $p(1-p)^3$ |
| $SFSS$ | 3 | $p^3(1-p)$ | $FFSS$ | 2 | $p^2(1-p)^2$ |
| $SFSF$ | 2 | $p^2(1-p)^2$ | $FFSF$ | 1 | $p(1-p)^3$ |
| $SFFS$ | 2 | $p^2(1-p)^2$ | $FFFS$ | 1 | $p(1-p)^3$ |
| $SFFF$ | 1 | $p(1-p)^3$ | $FFFF$ | 0 | $(1-p)^4$ |

$$b(3; 4, p) = P(FSSS) + P(SFSS) + P(SSFS) + P(SSSF)$$
$$= 4p^3(1-p)$$

# Theorem

- Since the ordering of *S*'s and *F*'s is not important, the second factor in the previous equation is $p^x * (1-p)^{n-x}$ (e.g., the first *x* trials resulting in *S* and the last resulting in *F*).
- The first factor is the number of ways of choosing *x* of the *n* trials to be *S*'s—that is, the number of combinations of size *x* that can be constructed from *n* distinct objects (trials here).

$${}^nC_x \, p^x (1-p)^{n-x}$$

$$b(x; n, p) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & x = 0, 1, 2, \ldots, n \\ 0 & \text{otherwise} \end{cases}$$

# Example

- Each of six randomly selected cola drinkers is given a glass containing cola *S* and one containing cola *F*. The glasses are identical in appearance except for a code on the bottom to identify the cola. Suppose there is actually no tendency among cola drinkers to prefer one cola to the other. Then p= P(a selected individual prefers S)=0.5. S0, with X= the number among the six who prefer *S*

$X = \{0, 1, 2, 3 \dots 6\}$

$P(X=3) = ?$

$p = 0.5, \quad n = 6$

$^nC_x \, p^x (1-p)^{n-x}$

$^6C_3 (0.5)^3 (1-0.5)^{6-3} = 20 \times (0.5)^6 = 0.3125$

# Exercise

- Suppose that 20% of all copies of a particular textbook fail a certain binding strength test. Let *X* denote the number among 15 randomly selected copies that fail the test. Then *X* has a binomial distribution with n=15 and p=.2
- What is the probability that at most 8 fail the test?

$$P(X \le 8) = \sum_{x=0}^{8} b(x, 15, 0.2)$$

$$= B(8, 15, 0.2)$$

cummulative prob.

from Binomial table got value

$$= 0.999$$

# Solution

- What is the probability that exactly 8 fail?

$P(X=8) = P(X \leq 8) - P(X \leq 7)$
(0 to 8) (0 to 7)
$= B(8,15,0.2) - B(7,15,0.2)$
$= 0.999 - 0.996 = 0.003$

- What is the probability that at least 8 fail?

$P(X \geq 8) = ? \quad 1 - P(X \leq 7)$
$= 1 - B(7,15,0.2) = 1 - 0.996$
$= 0.004$

- What is the probability that fail is between 4 and 7 (inclusive)

$P(X \leq 7) - P(X \leq 3)$
$= B(7,15,0.2) - B(3,15,0.2)$
$= 0.996 - 0.648$
$= 0.348$

# Definitions: Bernouilli

- **Bernouilli trial:** If there is only 1 trial with probability of success *p* and probability of failure *1-p*, this is called a Bernouilli distribution. (special case of the binomial with *n*=1)

- Probability of success:

$$P(X=1)=\binom{1}{1}p^{1}(1-p)^{1-1}=p$$

- Probability of failure:

$$P(X=0)=\binom{1}{0}p^{0}(1-p)^{1-0}=1-p$$

# Characteristics of Bernouilli distribution

innovate achieve lead

For Bernouilli ($n=1$)

$E(X) = p$

$Var\ (X) = p(1-p)$

# Expected value and variance of Binomial Distribution

If $X$ follows a binomial distribution with parameters $n$ and $p$:

$$X \sim Bin\,(n, p)$$

Then:

$\mu_X = E(X) = np$

$\sigma_X^2 = Var\,(X) = np(1-p)$ $= npq$ (here $q = 1-p$)

$\sigma_X = SD\,(X) = \sqrt{np(1-p)}$

# Variance Proof (optional!)

For Y~Bernouilli (p)

Y=1 if yes

Y=0 if no

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

For X~Bin (N,p)

$$X = \sum_{i=1}^{n} Y_{Bernouilli}; Var(Y) = p(1-p)$$

$$= Var(X) = Var(\sum_{i=1}^{n} Y) = \sum_{i=1}^{n} Var(Y) = np(1-p)$$

# Poisson distribution

Poisson distribution is for counts—if events happen at a constant rate over time, the Poisson distribution gives the probability of X number of events occurring in time T.

# Poisson Mean and Variance

Mean

$$\boxed{\mu = \lambda}$$ rate

**For a Poisson random variable, the variance and mean are the same!**

- Variance and Standard Deviation

$$\boxed{\sigma^2 = \lambda}$$

$$\boxed{\sigma = \sqrt{\lambda}}$$

where $\lambda$ = expected number of hits in a given time period

# Poisson Distribution, example

- The Poisson distribution models counts, such as the number of new cases of COVID that occur in women in Bangalore next month.
- The distribution tells you the probability of all possible numbers of new cases, from 0 to infinity.
- If X= # of new cases next month and $X \sim$ Poisson ($\lambda$), then the probability that $X = k$ (a particular count) is:

$$\boxed{p(X=k) = \frac{\lambda^k e^{-\lambda}}{k!}}$$

# Example: Poisson distribution

- Suppose that a rare disease has an incidence of 1 in 1000 person-years. Assuming that members of the population are affected independently, find the probability of k cases in a population of 10,000 (followed over 1 year) for k=0,1,2.
- The expected value (mean) =$\lambda$ = .001*10,000 = 10
- 10 new cases expected in this population per year→

$$\frac{\lambda^{x} e^{-\lambda}}{x!_0}$$

$$P(X=0)=\frac{(10)^0 e^{-(10)}}{0!}=.0000454$$

$$P(X=1)=\frac{(10)^1 e^{-(10)}}{1!}=.000454$$

$$P(X=2)=\frac{(10)^2 e^{-(10)}}{2!}=.00227$$

# more on Poisson…

- "Poisson Process" (rates)
- Note that the Poisson parameter $\lambda$ can be given as the mean number of events that occur in a defined time period OR, equivalently, $\lambda$ can be given as a rate, such as $\lambda$=2/month (2 events per 1 month) that must be multiplied by *t*=time (called a "Poisson Process") $\rightarrow$
- X ~ Poisson ($\lambda$)

$$P(X=k)=\frac{(\lambda t)^k e^{-\lambda t}}{k!}$$

E(X) = $\lambda$t
Var(X) = $\lambda$t

# Practice problems

1a. If the calls received on your mobile phone follow the Poisson distribution with a constant rate λ=4 calls per hour, what's the probability that, if you forget to turn your phone off in a 1.5 hour theater play, your phone rings during that time?

$\lambda = 4$ calls/hr

$\lambda t = 4 * 1.5 = 6$

$P(X \geq 1) = 1 - P(0) = 1 - \left[\frac{(6)^0 e^{-6}}{0!}\right] \approx 0.9$

1b. How many avg. phone calls do you expect to get during the play?

$E(X) = \lambda t = 4 \times 1.5 = 6$

# Poisson distribution as limit

- In any binomial experiment in which n is large and p is small, $b(x;n,p)$ is approximately equal to $p(x;\mu)$ , where $\mu = np$ .

- As a rule of thumb, this approximation can safely be applied if $n > 50$ and $np < 5$ .

# Example

- If a publisher of nontechnical books takes great pains to ensure that its books are free of typographical errors, so that the probability of any given page containing at least one such error is .005 and errors are independent from page to page, what is the probability that one of its 400-page novels will contain exactly one page with errors? At most three pages with errors?

**Solution**

- With *S* denoting a page containing at least one error and *F* an error-free page, the number *X* of pages containing at least one error is a binomial rv with n=400 and p=.005,so np=2.

# Solution

Binomial distribution yes $\quad n = 400 \quad p = 0.005$

$np = ? \quad 2$

$\because \; n > 50 \;\&\; np < 5$ we can apply poison dis.

$P(X=1)$ then $\mu = \lambda = np = 2$.

$\frac{2 \times e^{-2}}{1!} = 0.27$

$P(X \leq 3)$

$= \sum_{x=0}^{3} \frac{e^{-2} \times 2^{x}}{x!} = 0.857$

## Table 3.2 Comparing the Poisson and Three Binomial Distributions

| $x$ | $n = 30, p = .1$ | $n = 100, p = .03$ | $n = 300, p = .01$ | Poisson, $\mu = 3$ |
|---|---|---|---|---|
| 0 | 0.042391 | 0.047553 | 0.049041 | 0.049787 |
| 1 | 0.141304 | 0.147070 | 0.148609 | 0.149361 |
| 2 | 0.227656 | 0.225153 | 0.224414 | 0.224042 |
| 3 | 0.236088 | 0.227474 | 0.225170 | 0.224042 |
| 4 | 0.177066 | 0.170606 | 0.168877 | 0.168031 |
| 5 | 0.102305 | 0.101308 | 0.100985 | 0.100819 |
| 6 | 0.047363 | 0.049610 | 0.050153 | 0.050409 |
| 7 | 0.018043 | 0.020604 | 0.021277 | 0.021604 |
| 8 | 0.005764 | 0.007408 | 0.007871 | 0.008102 |
| 9 | 0.001565 | 0.002342 | 0.002580 | 0.002701 |
| 10 | 0.000365 | 0.000659 | 0.000758 | 0.000810 |

# The Mean and Variance of X

- Since as $b(x;n,p) \rightarrow p(x;\mu)$ as $n \rightarrow \infty$, $p \rightarrow 0$, $np \rightarrow \mu$, the mean and variance of a binomial variable should approach those of a Poisson variable.

- These limits are $np \rightarrow \mu$ and $np(1 - p) \rightarrow \mu$

If $X$ has a Poisson distribution with parameter $\mu$, then $E(X) = V(X) = \mu$.

# Normal distribution

- symmetric bell shape
- mean and median are equal; both located at the center of the distribution

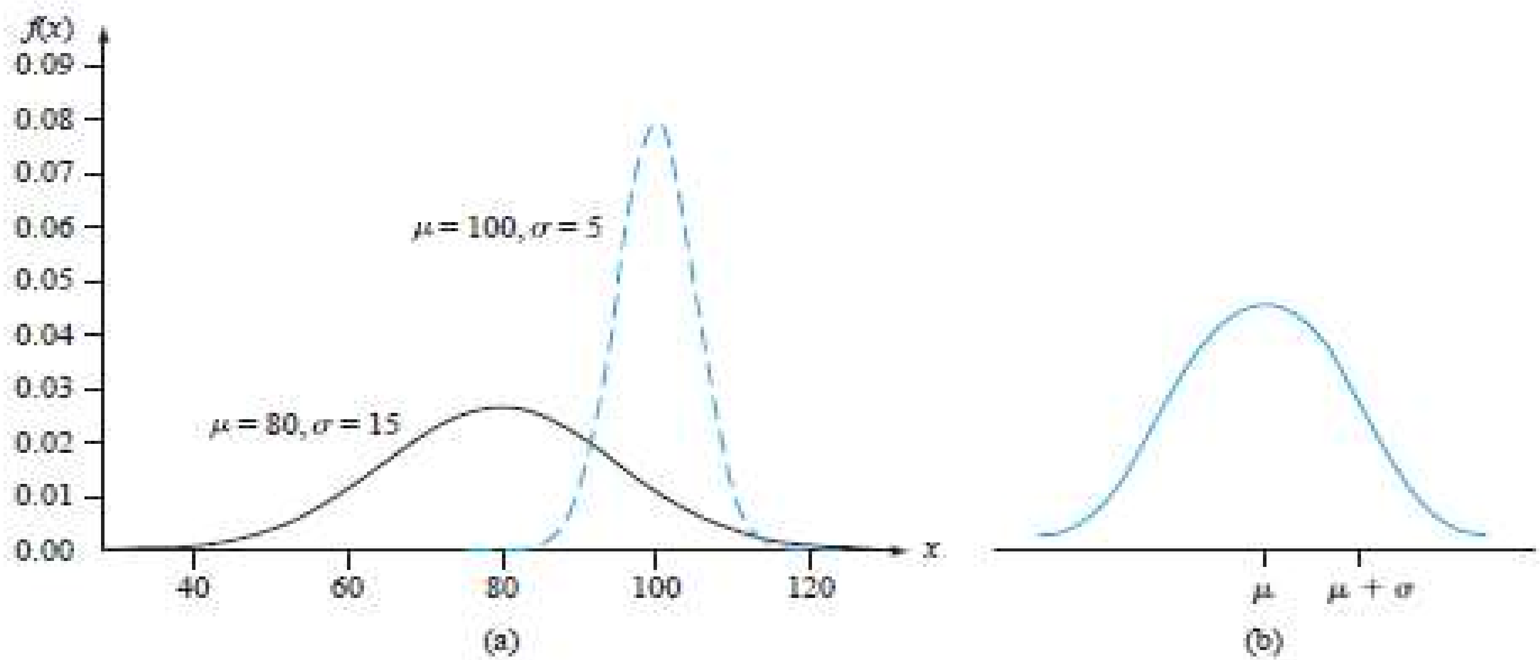



**Figure 4.13** (a) Two different normal density curves (b) Visualizing $\mu$ and $\sigma$ for a normal distribution

# Observations of Normal Distributions

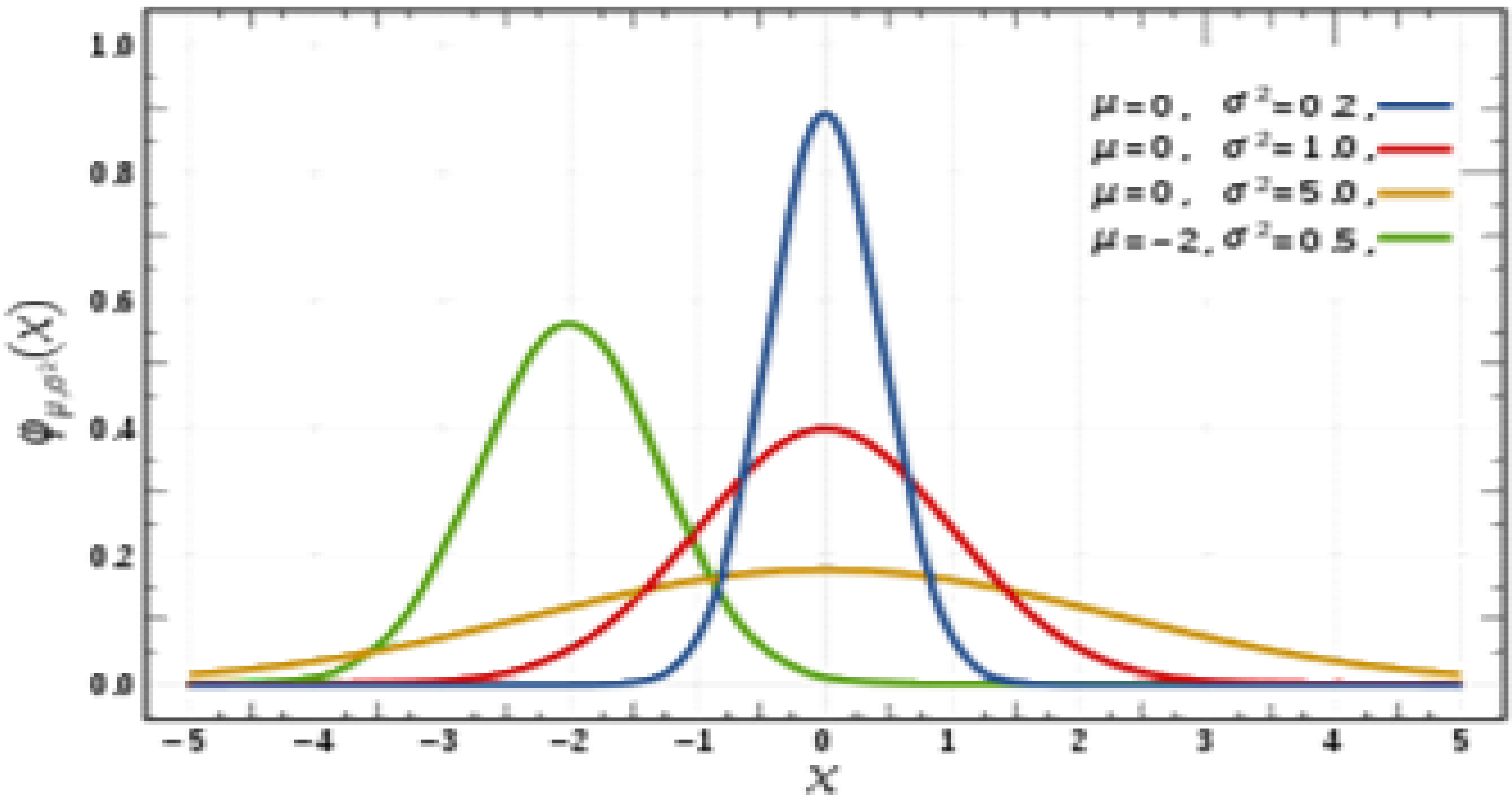



Image: google

# Standard Deviation Rule for Normal Random Variables

In general, if X is a normal random variable, then the probability is

- 68% that X falls within 1 σ of μ , that is, in the interval **μ±σ**
- 95% that X falls within 2 σ of μ , that is, in the interval **μ±2σ**
- 99.7% that X falls within 3 σ of μ , that is, in the interval **μ±3σ**

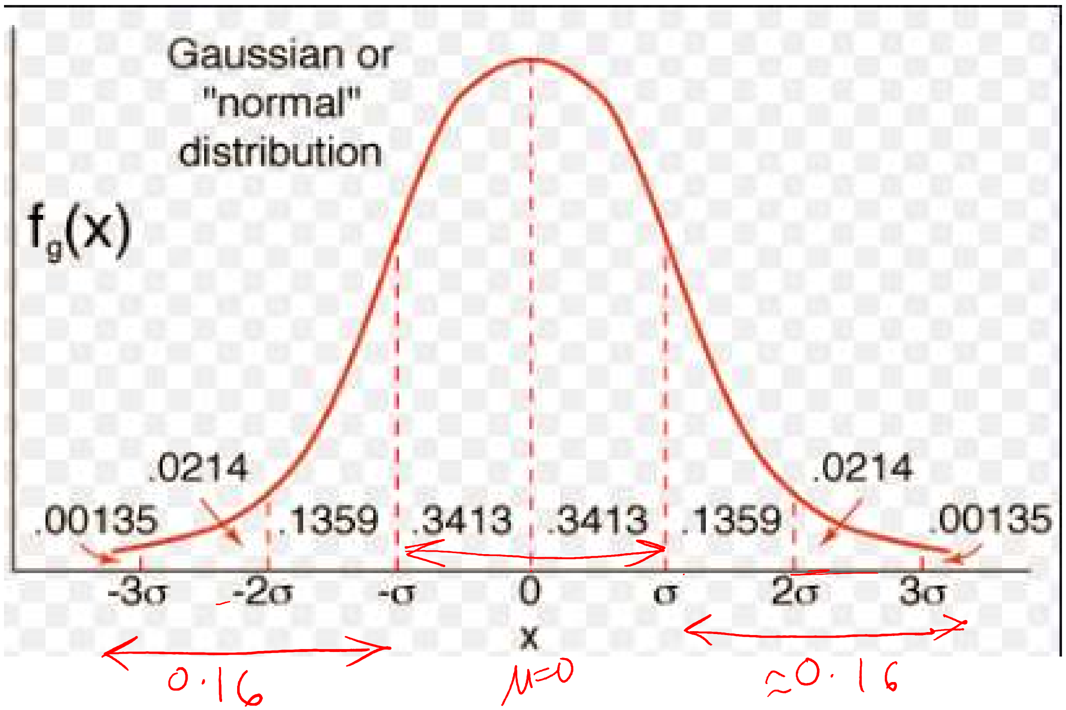
Using probability notation, we may write

$$0.68 = P(\mu - \sigma < X < \mu + \sigma)$$

$$0.95 = P(\mu - 2\sigma < X < \mu + 2\sigma)$$

$$0.997 = P(\mu - 3\sigma < X < \mu + 3\sigma)$$

# Normal Distribution




- since 0.68 is the probability of being within 1 standard deviation of the mean,
- (1 - .68) / 2 = 0.16 is the probability of being further than 1 standard deviation below the mean (or further than 1 standard deviation above the mean).
- Likewise, (1 - .95) / 2 = 0.025 is the probability of being more than 2 standard deviations below (or above) the mean;
- (1 - .997) / 2 = 0.0015 is the probability of being more than 3 standard deviations below (or above) the mean.

# Exercise

- **Suppose that hair length of a randomly chosen female is a normal random variable with mean μ=11 and standard deviation σ=1.5.**

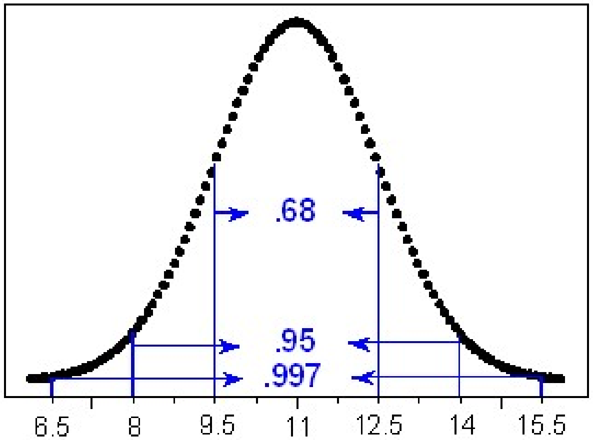
Ques1. What is the probability that a randomly chosen female will have hair length between 8 and 14 inches? $11 \pm 2*1.5$ is 8 to 14 range

$\mu \pm 2\sigma \rightarrow$ prob? $\rightarrow 0.95$

Ques2. A female is almost guaranteed (.997 probability) to have hair length between what two values?

$\mu \pm 3\sigma = 11 \pm 2*1.5$ i.e. 6.5 to 15.5

Ques3. The probability is only 2.5% female will have hair length greater than how many inches?




95% of people/female have value b/w 8 to 14 inch.

2.5 of female have < 8 inch length & 2.5 have > 14 inch length.

# The Normal Distribution

- The normal distribution is the most important one in all of probability and statistics.
- Many numerical populations have distributions that can be fit very closely by an appropriate normal curve.
- Examples include heights, weights, and other physical characteristics

A continuous rv $X$ is said to have a **normal distribution** with parameters $\mu$ and $\sigma$ (or $\mu$ and $\sigma^2$), where $-\infty < \mu < \infty$ and $0 < \sigma$, if the pdf of $X$ is

$$f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/(2\sigma^2)} \qquad -\infty < x < \infty \qquad (4.3)$$

# Standard Normal Distribution

- The normal distribution with parameter values $\mu=0$ and $\sigma=1$ is called the **standard normal distribution.**

# Non-standard Normal Distributions

- Every unique pair of μ and σ values defines a different normal distribution

- Fortunately, a mechanism was developed by which all normal distributions can be converted into a single distribution: the z distribution.

- This process yields the **standardized normal distribution** (or curve).

$$z = \frac{x - \mu}{\sigma}, \quad \sigma \neq 0$$

# Probabilities using z value

If $X$ has a normal distribution with mean $\mu$ and standard deviation $\sigma$, then

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

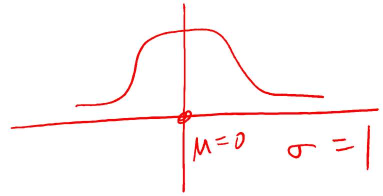
has a standard normal distribution. Thus

$$P(a \le X \le b) = P\left(\frac{a - \mu}{\sigma} \le Z \le \frac{b - \mu}{\sigma}\right)$$

$$= \Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right)$$

$$P(X \le a) = \Phi\left(\frac{a - \mu}{\sigma}\right) \qquad P(X \ge b) = 1 - \Phi\left(\frac{b - \mu}{\sigma}\right)$$

# z - score




- A z score is the number of standard deviations that a value, x, is above or below the mean.

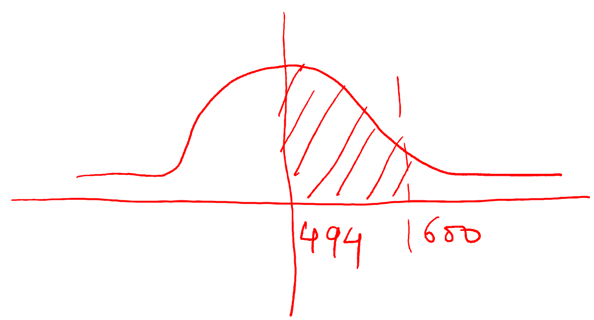
- If the value of x is less than the mean, the z score is negative;

- If the value of x is more than the mean, the z score is positive; and

- If the value of x equals the mean, the associated z score is zero.

# Example

The Graduate Management Aptitude Test (GMAT), produced by the Educational Testing Service in Princeton, New Jersey, is widely used by graduate schools of business in the United States as an entrance requirement. Assuming that the scores are normally distributed, probabilities of achieving scores over various ranges of the GMAT can be determined. In a recent year, the mean GMAT score was 494 and the standard deviation was about 100. What is the probability that a randomly selected score from this administration of the GMAT is between 600 and the mean?

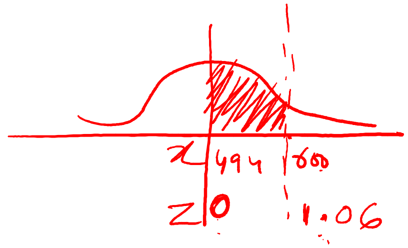


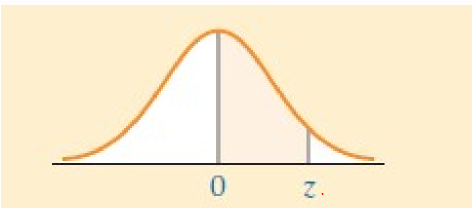
# Solution

for $x = 600$, $Z = \frac{x-\mu}{\sigma} = \frac{600-494}{100} = 1.06$

$P(494 \le x \le 600 \mid \mu = 494 \text{ and } \sigma = 100) = ?$

area under curve or prob $= 0.3554$

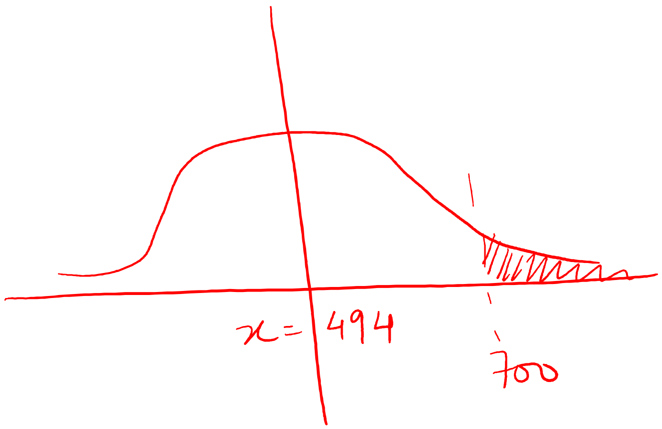



| SECOND DECIMAL PLACE IN $z$ | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $z$ | 0.00 | 0.01 | 0.02 | 0.03 | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 |
| 0.0 | .0000 | .0040 | .0080 | .0120 | .0160 | .0199 | .0239 | .0279 | .0319 | .0359 |
| 0.1 | .0398 | .0438 | .0478 | .0517 | .0557 | .0596 | .0636 | .0675 | .0714 | .0753 |
| 0.2 | .0793 | .0832 | .0871 | .0910 | .0948 | .0987 | .1026 | .1064 | .1103 | .1141 |
| 0.3 | .1179 | .1217 | .1255 | .1293 | .1331 | .1368 | .1406 | .1443 | .1480 | .1517 |
| 0.4 | .1554 | .1591 | .1628 | .1664 | .1700 | .1736 | .1772 | .1808 | .1844 | .1879 |
| 0.5 | .1915 | .1950 | .1985 | .2019 | .2054 | .2088 | .2123 | .2157 | .2190 | .2224 |
| 0.6 | .2257 | .2291 | .2324 | .2357 | .2389 | .2422 | .2454 | .2486 | .2517 | .2549 |
| 0.7 | .2580 | .2611 | .2642 | .2673 | .2704 | .2734 | .2764 | .2794 | .2823 | .2852 |
| 0.8 | .2881 | .2910 | .2939 | .2967 | .2995 | .3023 | .3051 | .3078 | .3106 | .3133 |
| 0.9 | .3159 | .3186 | .3212 | .3238 | .3264 | .3289 | .3315 | .3340 | .3365 | .3389 |
| 1.0 | .3413 | .3438 | .3461 | .3485 | .3508 | .3531 | .3554 | .3577 | .3599 | .3621 |

# Exercise

What is the probability of obtaining a score greater than 700 on a GMAT test that has a mean of 494 and a standard deviation of 100? Assume GMAT scores are normally distributed.

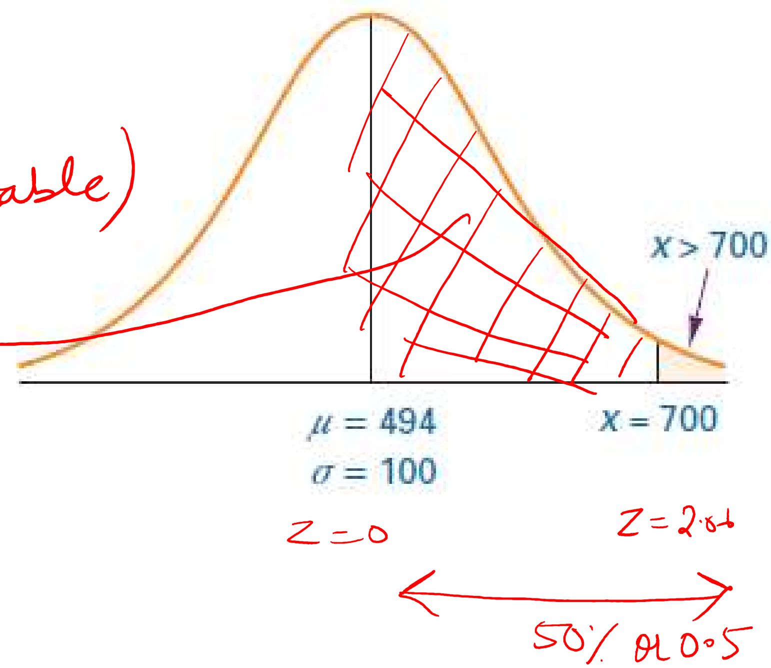


# Solution

for $x = 700$, $z = \dfrac{x-\mu}{\sigma} = \dfrac{700-494}{100} = 2.06$

for $z = 2.06$

prob. = ? $= 0.4803$ (from z table)




Ans. $P(X > 700)$

$= 0.5 - 0.4803$

$= 0.0197$

# Exercise

For the same GMAT examination, what is the probability of randomly drawing a score that is 550 or less?

$\mu = 494$

$\sigma = 100$

$P(x \le 550) = ?$

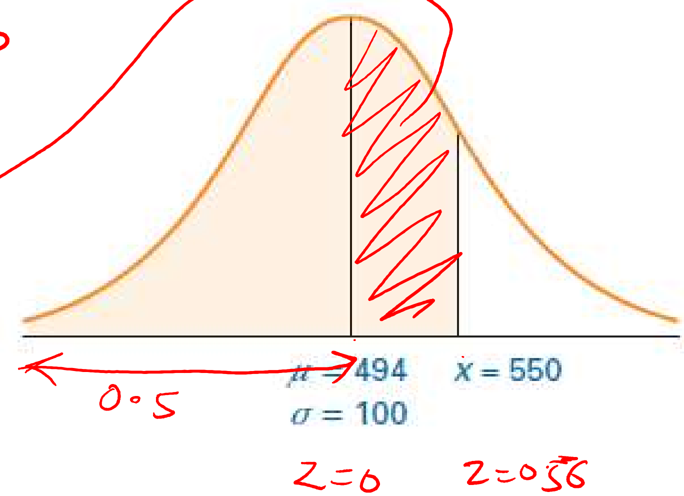
# Solution

for $x = 550$

$$Z = \frac{550 - 494}{100} = 0.56$$

for $z = 0.56$

prob $= 0.2123$




$$P(X \leq 550) = 0.5 + 0.2123$$

$$= 0.7123$$

# Exercise

What is the probability of randomly obtaining a score between 300 and 600 on the GMAT exam?

$$P(300 \leq X \leq 600)$$

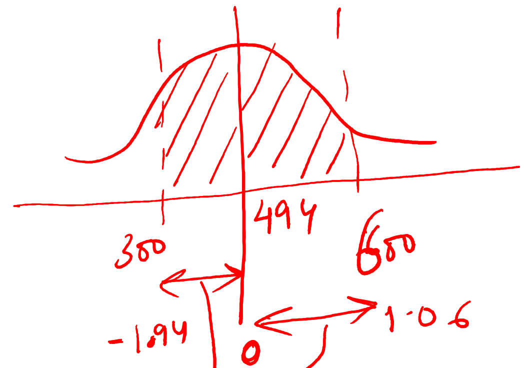
$$\mu = 494$$

$$\sigma = 100$$

# Solution

for $x = 300$ $\quad z = ? \quad = \frac{300-494}{100} = -1.94$

for $x = 600$ $\quad z = \frac{x-\mu}{\sigma} = \frac{600-494}{100} = 1.06$




for $z = 1.06$ prob. is 0.3554

for $z = -1.94$ " is 0.4738

$$\text{Ans} = \begin{array}{r} 0.3554 \\ +\ 0.4738 \\ \hline 0.8292 \end{array}$$

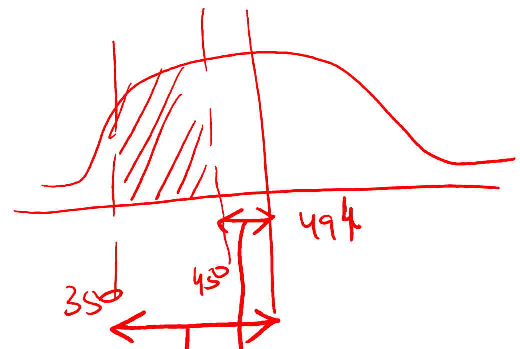
# Exercise (HW)

What is the probability of getting a score between 350 and 450 on the same GMAT exam?

# Solution

for $x = 350$, $z = \frac{350 - 494}{100} = -1.44$

for $x = 450$, $z = \frac{450 - 494}{100} = -0.44$




prob. for $z = -1.44$ is $0.4251$

prob. for $z = -0.44$ is $0.1700$

$$
\text{Ans} = \begin{array}{r} 0.4251 \\ -\;0.1700 \\ \hline 0.2551 \\ \hline \end{array}
$$

# Exercise

Runzheimer International publishes business travel costs for various cities throughout the world. In particular, they publish per diem totals, which represent the average costs for the typical business traveler including three meals a day in business-class restaurants and single-rate lodging in business-class hotels and motels. If 86.65% of the per diem costs in Buenos Aires, Argentina, are less than $449 and if the standard deviation of per diem costs is $36, what is the average per diem cost in Buenos Aires? Assume that per diem costs are normally distributed.

86.65% values below 449 $


$z = ?$

$0.8665 - 0.5 = 0.3665$

# Solution

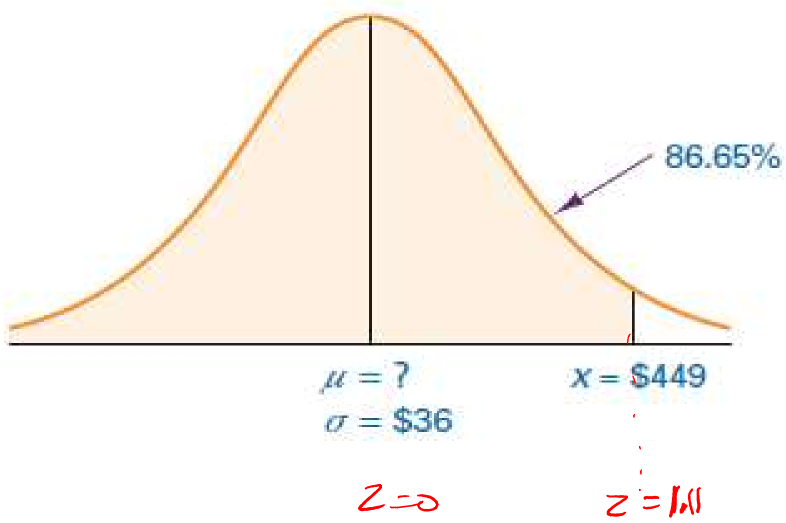
for area = 0.3665

z value is 1.11

$$Z = \frac{x - \mu}{\sigma}$$

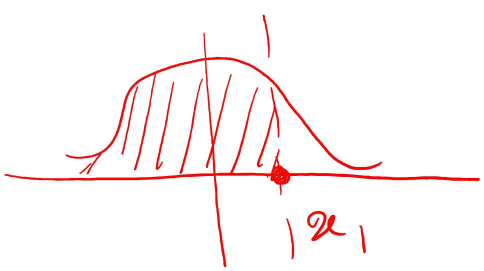
$$1.11 = \frac{449 - \mu}{36}$$

$$\mu = 409.04\,\$$$



# Exercise (HW)

The U.S. Environmental Protection Agency publishes figures on solid waste generation in the United States. One year, the average number of waste generated per person per day was 3.58 pounds. Suppose the daily amount of waste generated per person is normally distributed, with a standard deviation of 1.04 pounds. Of the daily amounts of waste generated per person, 67.72% would be greater than what amount?




$x < x_1$


$x > x_2$

$0.6772 - 0.5$

$= 0.1772$

↑ area

$z = ?$

# Solution

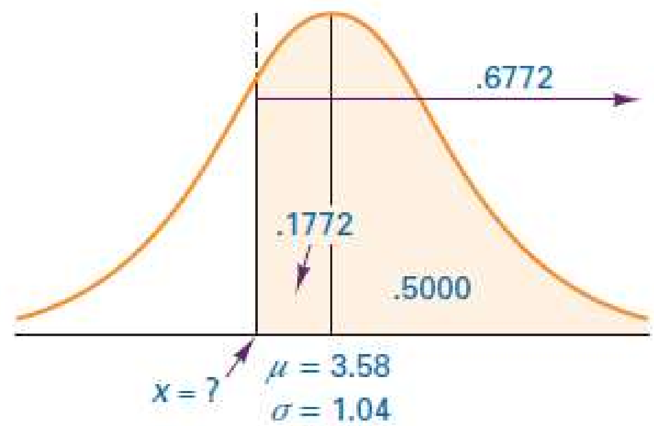
area = 0.1772

↓

z = 0.46

left side of curve

z will -ve.




$$-0.46 = \frac{x - 3.58}{1.04}$$

$$\underline{x = 3.10}$$