Class 11

Daniel B. Rowe, Ph.D.

Department of Mathematical and Statistical Sciences

Agenda:

Recap Chapter 6.1 – 6.5

Lecture Chapter 7.2 – 7.3

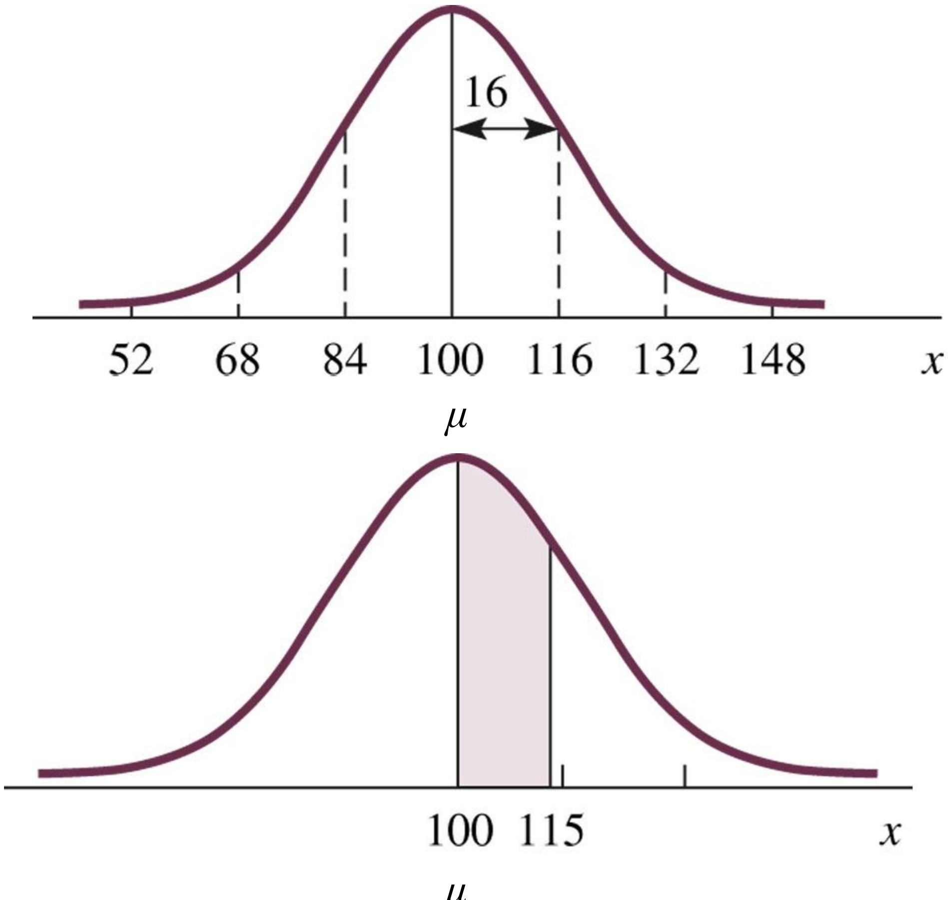
Recap Chapter 6.1 - 6.5

6: Normal Probability Distributions

6.3 Applications of Normal Distributions

Example:
Assume that IQ scores are normally distributed with a mean μ of 100 and a standard deviation σ of 16.

If a person is picked at random, what is the probability that his or her IQ is between 100 and 115?
i.e. $P(100 < x < 115)$?



Figures from Johnson & Kuby, 2012.

6: Normal Probability Distributions

6.3 Applications of Normal Distributions

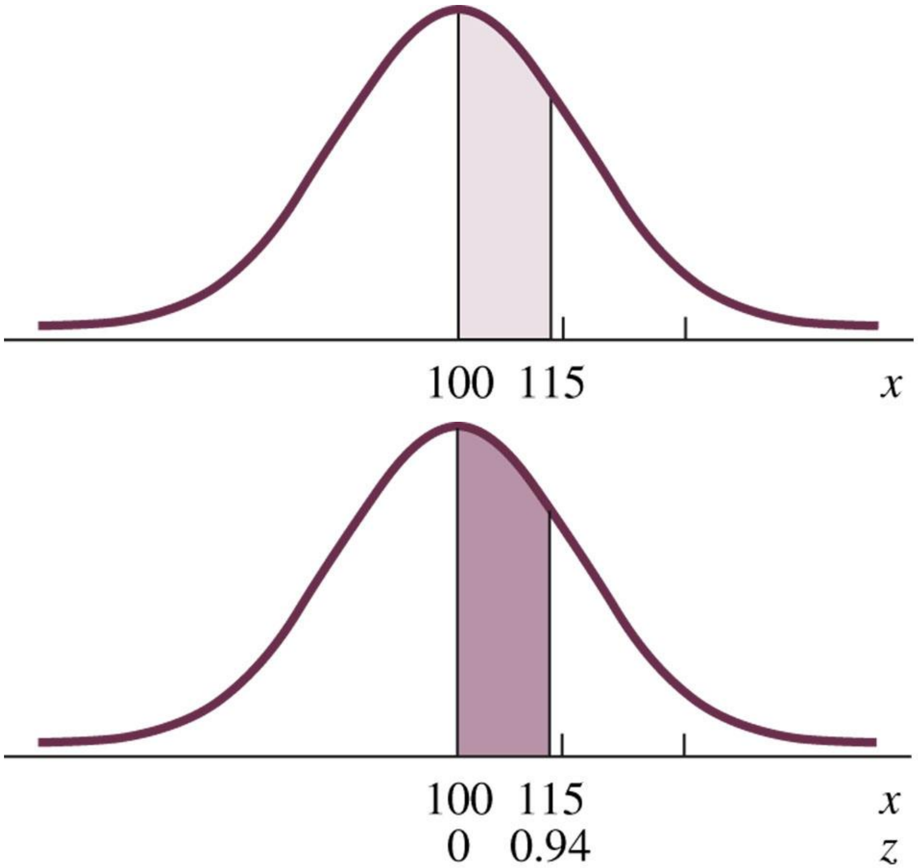
IQ scores normally distributed
μ=100 and σ=16.

$$P(100 < x < 115)$$

$$z = \frac{x - \mu}{\sigma} \qquad \begin{aligned} x_1 &= 100 \\ x_2 &= 115 \end{aligned}$$

$$z_1 = \frac{x_1 - \mu}{\sigma} = \frac{100 - 100}{16} = 0$$

$$z_2 = \frac{x_2 - \mu}{\sigma} = \frac{115 - 100}{16} = 0.94$$

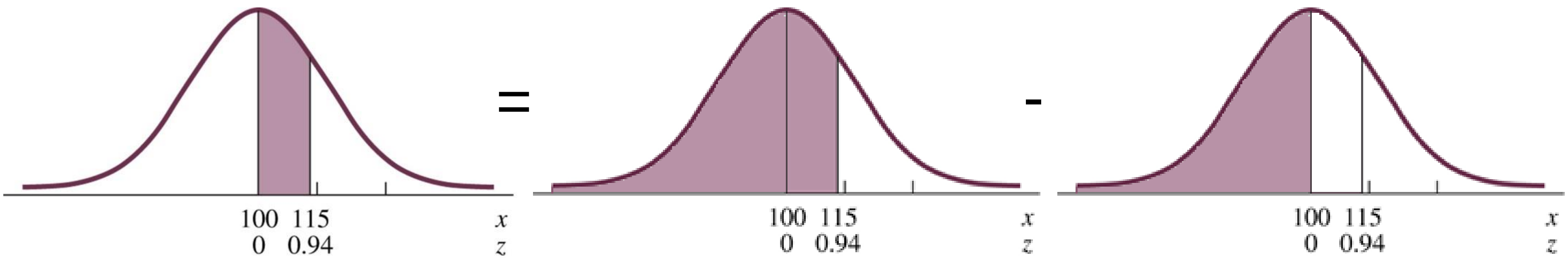


Figures from Johnson & Kuby, 2012.

6: Normal Probability Distributions

6.3 Applications of Normal Distributions

Now we can use the table.



$$\begin{aligned} P(0 < z < 0.94) &= P(z < 0.94) - P(z < 0) \\ &= 0.8264 - .5 \\ &= 0.3264 \end{aligned}$$

Figures from Johnson & Kuby, 2012.

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389

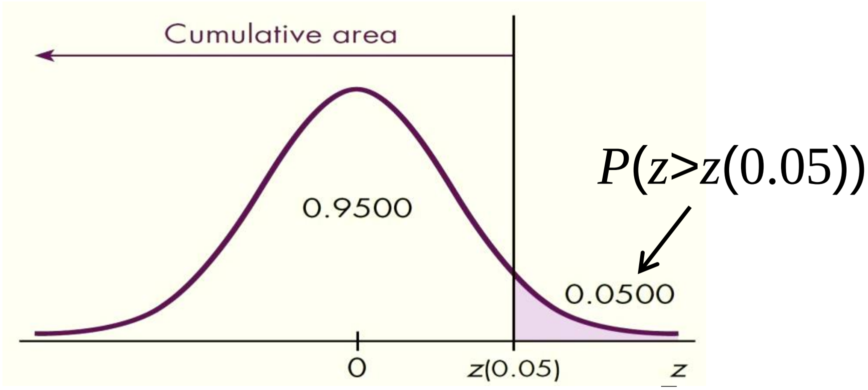
6: Normal Probability Distributions

6.4 Notation

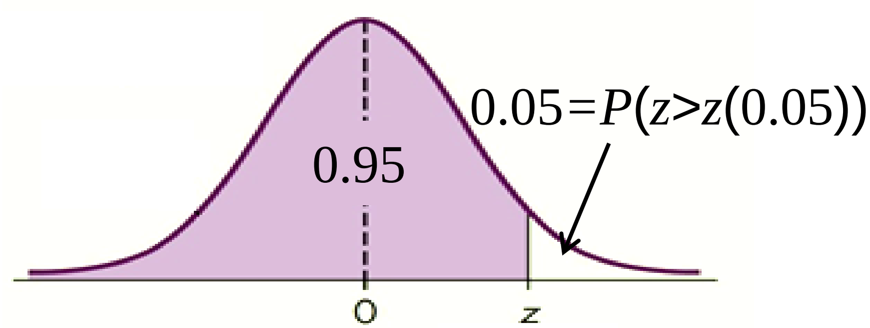
Example:

Let α=0.05. Let's find $z(0.05)$.

$P(z>z(0.05))=0.05$.



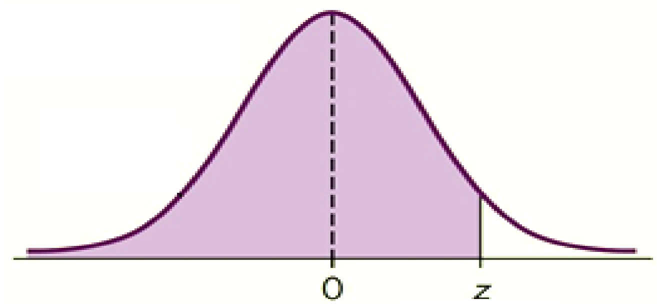
Same as finding $P(z<z(0.05))=1-0.05$.



Figures from Johnson & Kuby, 2012.

6: Normal Probability Distributions

6.4 Notation



Example:

Same as finding $P(z<z(0.05))=0.95$.

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5754
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7258	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7518	0.7549
0.7	0.7580	0.7612	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7996	0.8023	0.8051	0.8079	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
⋮										
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9430	0.9441
1.6	0.9452	0.9463	0.9474	0.9485	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9700	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9762	0.9767
2.0	0.9773	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9865	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936

1.645

Figures from Johnson & Kuby, 2012.

6: Normal Probability Distributions

n=14, p=1/2

6.5 Normal Approximation of the Binomial Distribution

From the binomial formula

$$P(4)=\frac{14!}{4!(14-4)!}(.5)^4(1-.5)^{14-4}$$

$$P(x=4)=0.061$$

From the Normal Distribution

$$P(3.5<x<4.5) \qquad \mu=7,\ \sigma^2=3.5$$

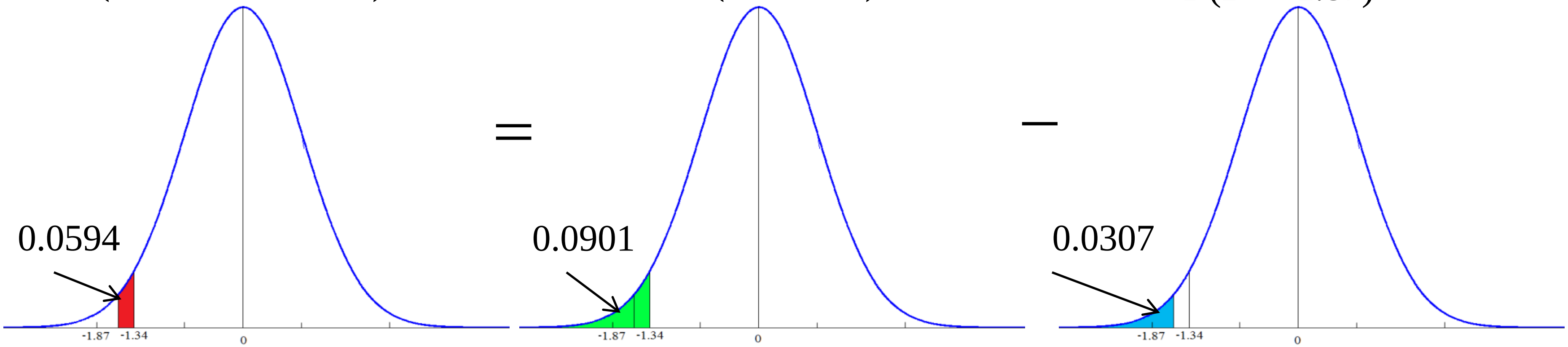
$$\sigma=1.87$$

$$z_1=\frac{x_1-\mu}{\sigma}=\frac{3.5-7}{1.87}=\text{-}1.87$$

$$z_2=\frac{x_2-\mu}{\sigma}=\frac{4.5-7}{1.87}=-1.34$$

$$P(-1.87<z<1.34)=0.0594$$

$P(-1.87<z<-1.34)$ = $P(z<-1.34)$ − $P(z<-1.87)$



6: Normal Probability Distributions

Questions?

Homework: Read Chapter 6.1-6.2

Web Assign

Chapter 6 # 7a&b, 9a&b, 13a, 19, 29, 31, 33, 41, 45, 47, 53, 61, 75, 95, 99

Not homework, but maybe fun to watch: NETFLIX

Lecture Chapter 7.2- 7.3

Chapter 7: Sample Variability

Daniel B. Rowe, Ph.D.

Department of Mathematical and Statistical Sciences

MARQUETTE UNIVERSITY

Be The Difference.

7: Sample Variability

7.2 The Sampling Distribution of Sample Means

When we take a random sample $x_1,\ldots, x_n$ from a population, one of the things that we do is compute the sample mean $\bar{x}$.

The value of $\bar{x}$ is not μ. Each time we take a random sample of size n, we get a different set of values $x_1,\ldots, x_n$ and a different value for $\bar{x}$.

7: Sample Variability

7.2 The Sampling Distribution of Sample Means

Recall: When we take a sample of data $x_1,\ldots,x_n$ from a population, then compute an estimate of a parameter it is called a sample statistic. i.e. $\bar{x}$ for μ

Sampling Distribution of a sample statistic: The distribution of values for a sample statistic obtained from repeated samples, all of the same size and all drawn from the same population.

7: Sample Variability

7.2 The Sampling Distribution of Sample Means

Let's discuss the relationship between the sample mean and the population mean.

Assume that we have a population of items with population mean μ and population standard deviation σ.

If we take a random sample of size n and compute sample mean, $\bar{x}$.

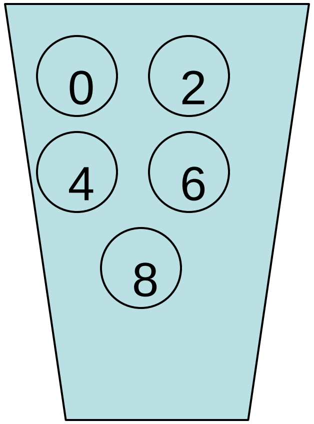
The collection of all possible means is called the *sampling distribution of the sample mean*.

7: Sample Variability

7.2 The Sampling Distribution of Sample Means

Example:

N=5 balls in bucket, select n=1 *with* replacement.



S={ }

Prob. of each value =

7: Sample Variability

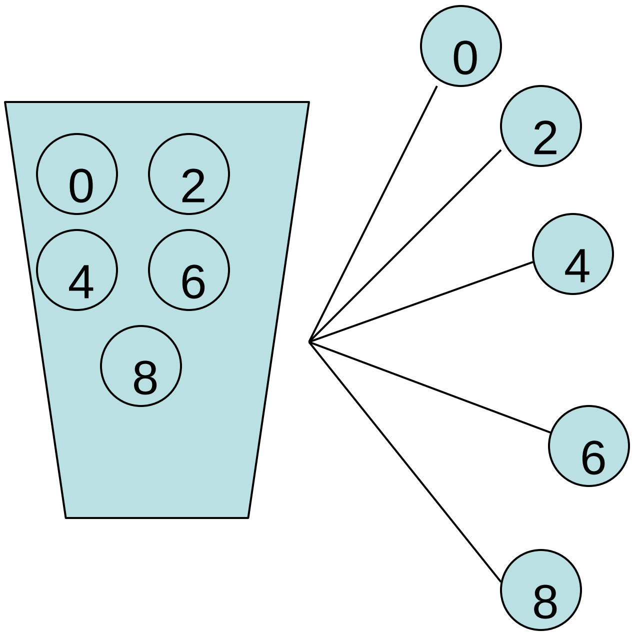
7.2 The Sampling Distribution of Sample Means

Example:

N=5 balls in bucket, select n=1 *with* replacement.

Population data values:
0, 2, 4, 6, 8.

5 possible values



S={0, 2, 4, 6, 8}

$x = 0,$ occurs one time

$x = 2,$ occurs one time

$x = 4,$ occurs one time

$x = 6,$ occurs one time

$x = 8,$ occurs one time

Prob. of each value = 1/5 = 0.2

7: Sample Variability

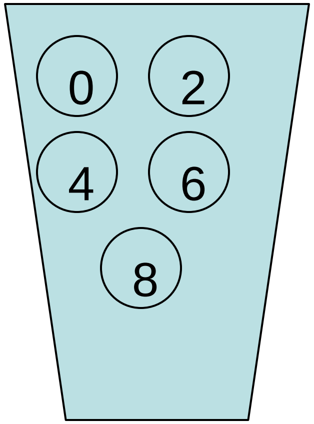
7.2 The Sampling Distribution of Sample Means

Example:

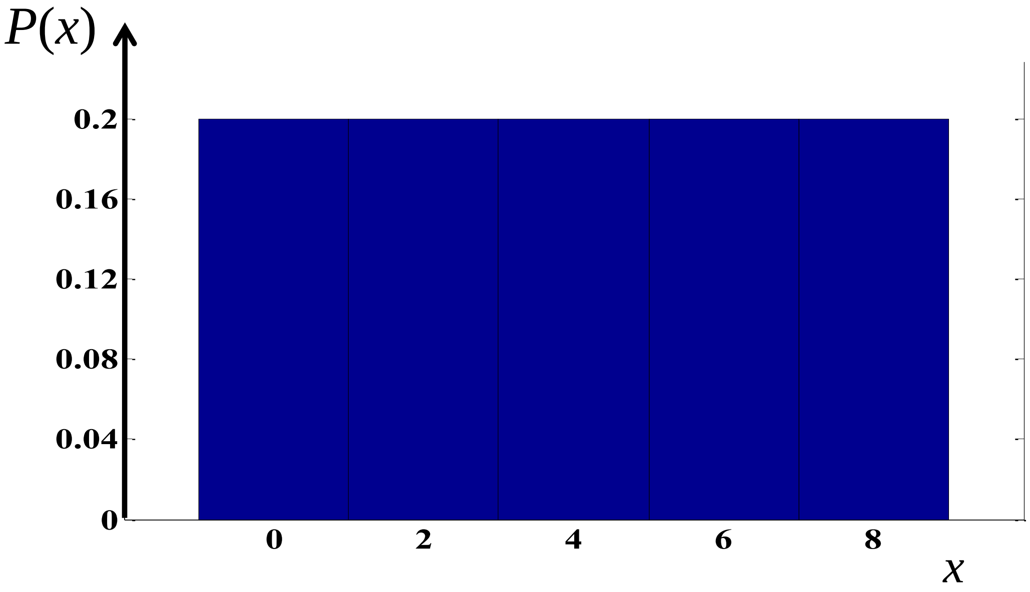
N=5 balls in bucket, select n=1 *with* replacement.

Population data values:
0, 2, 4, 6, 8.

5 possible values



x	$P(x)$
0	1/5
2	1/5
4	1/5
6	1/5
8	1/5



7: Sample Variability

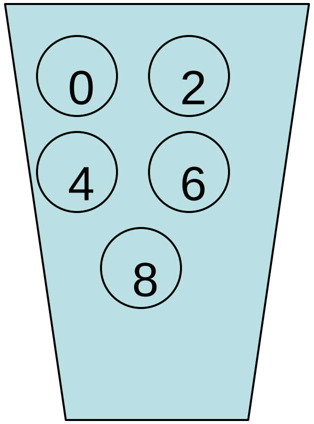
7.2 The Sampling Distribution of Sample Means

Example:

N=5 balls in bucket, select n=1 *with* replacement.

Population data values:
0, 2, 4, 6, 8.

5 possible values



x	$P(x)$
0	1/5
2	1/5
4	1/5
6	1/5
8	1/5

$$\mu = \sum [xP(x)]$$

$$=$$

7: Sample Variability

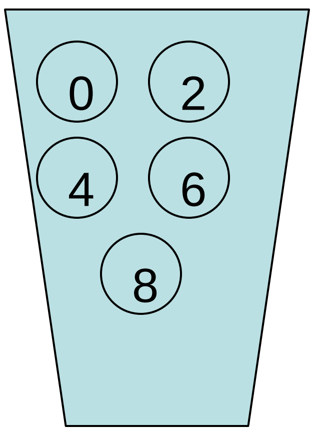
7.2 The Sampling Distribution of Sample Means

Example:

N=5 balls in bucket, select n=1 *with* replacement.

Population data values:
0, 2, 4, 6, 8.

5 possible values



x	$P(x)$
0	$1/5$
2	$1/5$
4	$1/5$
6	$1/5$
8	$1/5$

$$\mu = \sum [xP(x)]$$

$$= 0(1/5) + 2(1/5) + 4(1/5)$$

$$+ 6(1/5) + 8(1/5)$$

$$= 4$$

7: Sample Variability

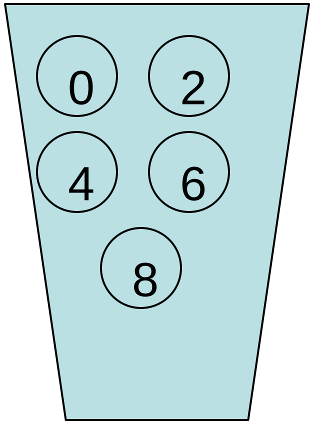
7.2 The Sampling Distribution of Sample Means

Example:

N=5 balls in bucket, select n=1 *with* replacement.

Population data values:
0, 2, 4, 6, 8.

5 possible values



x	$P(x)$
0	1/5
2	1/5
4	1/5
6	1/5
8	1/5

$$\sigma^2 = \sum[(x-\mu)^2 P(x)]$$

$$=$$

7: Sample Variability

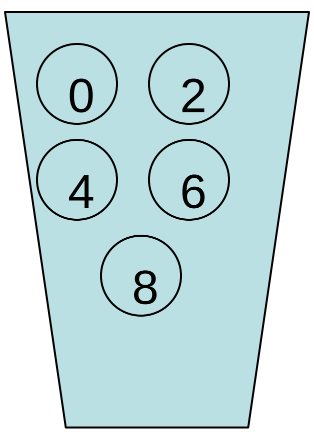
7.2 The Sampling Distribution of Sample Means

Example:

N=5 balls in bucket, select n=1 *with* replacement.

Population data values:
0, 2, 4, 6, 8.

5 possible values



x	$P(x)$
0	$1/5$
2	$1/5$
4	$1/5$
6	$1/5$
8	$1/5$

$$\sigma^2 = \sum[(x-\mu)^2 P(x)]$$

$$= (0-4)^2(1/5) + (2-4)^2(1/5)$$

$$= +(4-4)^2(1/5) + (6-4)^2(1/5)$$

$$+(8-4)^2(1/5)$$

$$= 8 \rightarrow \sigma = \sqrt{8} = 2\sqrt{2}$$

7: Sample Variability

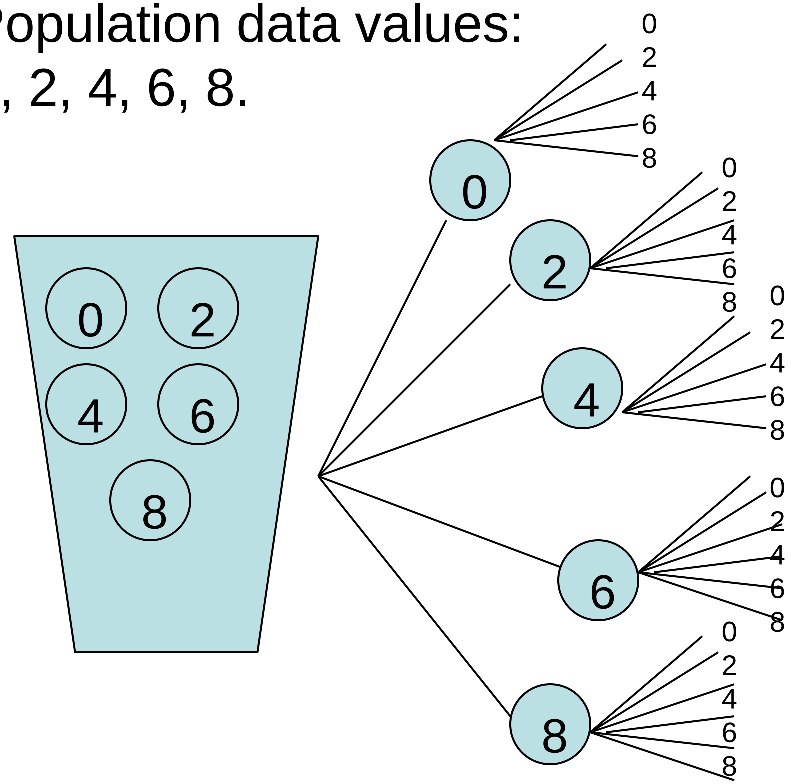
7.2 The Sampling Distribution of Sample Means

Example:

N=5 balls in bucket, select n=2 *with* replacement.

Population data values:

0, 2, 4, 6, 8.

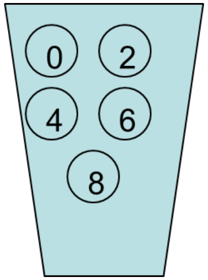


(0,0)	(2,0)	(4,0)	(6,0)	(8,0)
(0,2)	(2,2)	(4,2)	(6,2)	(8,2)
(0,4)	(2,4)	(4,4)	(6,4)	(8,4)
(0,6)	(2,6)	(4,6)	(6,6)	(8,6)
(0,8)	(2,8)	(4,8)	(6,8)	(8,8)

25 possible samples

7: Sample Variability

7.2 The Sampling Distribution of Sample Means



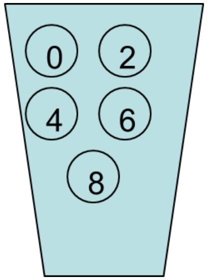
Example: There are N=5 items in the population.
Population data values: 0, 2, 4, 6, 8.
Take samples of size n=2 (with replacement).

There are 25 possible samples. Each sample has mean $\bar{x}$.

$(0,0)$	$(2,0)$	$(4,0)$	$(6,0)$	$(8,0)$	?	?	?	?	?
$(0,2)$	$(2,2)$	$(4,2)$	$(6,2)$	$(8,2)$	?	?	?	?	?
$(0,4)$	$(2,4)$	$(4,4)$	$(6,4)$	$(8,4)$	?	?	?	?	?
$(0,6)$	$(2,6)$	$(4,6)$	$(6,6)$	$(8,6)$	?	?	?	?	?
$(0,8)$	$(2,8)$	$(4,8)$	$(6,8)$	$(8,8)$	?	?	?	?	?

7: Sample Variability

7.2 The Sampling Distribution of Sample Means



Example: There are N=5 items in the population.
Population data values: 0, 2, 4, 6, 8.
Take samples of size n=2 (with replacement).

There are 25 possible samples.

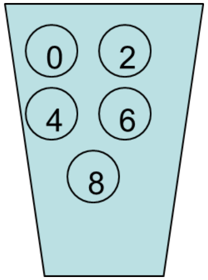
$(0,0)$	$(2,0)$	$(4,0)$	$(6,0)$	$(8,0)$
$(0,2)$	$(2,2)$	$(4,2)$	$(6,2)$	$(8,2)$
$(0,4)$	$(2,4)$	$(4,4)$	$(6,4)$	$(8,4)$
$(0,6)$	$(2,6)$	$(4,6)$	$(6,6)$	$(8,6)$
$(0,8)$	$(2,8)$	$(4,8)$	$(6,8)$	$(8,8)$

Each sample has mean $\bar{x}$.

0	1	2	3	4
1	2	3	4	5
2	3	4	5	6
3	4	5	6	7
4	5	6	7	8

7: Sample Variability

7.2 The Sampling Distribution of Sample Means



Example: N=5, values: 0, 2, 4, 6, 8, n=2 (with replacement).
25 possible samples.

Each possible sample is equally likely.

Prob. of each sample
= 1/25 = 0.04

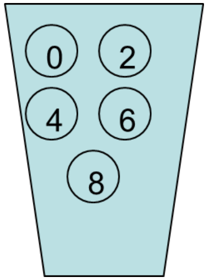
$$P[(i,j)] = 1/25$$
$$i = 0,2,4,6,8$$
$$j = 0,2,4,6,8$$

There are 25 possible samples.

$(0,0)$	$(2,0)$	$(4,0)$	$(6,0)$	$(8,0)$
$(0,2)$	$(2,2)$	$(4,2)$	$(6,2)$	$(8,2)$
$(0,4)$	$(2,4)$	$(4,4)$	$(6,4)$	$(8,4)$
$(0,6)$	$(2,6)$	$(4,6)$	$(6,6)$	$(8,6)$
$(0,8)$	$(2,8)$	$(4,8)$	$(6,8)$	$(8,8)$

7: Sample Variability

7.2 The Sampling Distribution of Sample Means



Example: N=5, values: 0, 2, 4, 6, 8, n=2 (with replacement).
25 possible samples. Each possible sample is equally likely.

Prob. of each samples mean = 1/25 = 0.04

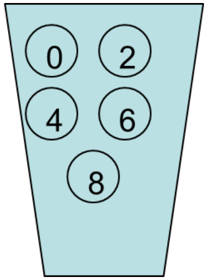
?	?	?	?	?
?	?	?	?	?
?	?	?	?	?
?	?	?	?	?
?	?	?	?	?

There are 25 possible samples.

(0,0)	(2,0)	(4,0)	(6,0)	(8,0)
(0,2)	(2,2)	(4,2)	(6,2)	(8,2)
(0,4)	(2,4)	(4,4)	(6,4)	(8,4)
(0,6)	(2,6)	(4,6)	(6,6)	(8,6)
(0,8)	(2,8)	(4,8)	(6,8)	(8,8)

7: Sample Variability

7.2 The Sampling Distribution of Sample Means



Example: N=5, values: 0, 2, 4, 6, 8, n=2 (with replacement).
25 possible samples. Each possible sample is equally likely.

Prob. of each samples mean = 1/25 = 0.04

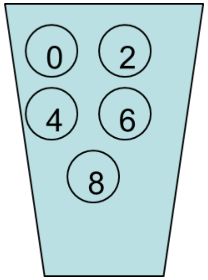
0	1	2	3	4
1	2	3	4	5
2	3	4	5	6
3	4	5	6	7
4	5	6	7	8

There are 25 possible samples.

(0,0)	(2,0)	(4,0)	(6,0)	(8,0)
(0,2)	(2,2)	(4,2)	(6,2)	(8,2)
(0,4)	(2,4)	(4,4)	(6,4)	(8,4)
(0,6)	(2,6)	(4,6)	(6,6)	(8,6)
(0,8)	(2,8)	(4,8)	(6,8)	(8,8)

7: Sample Variability

7.2 The Sampling Distribution of Sample Means



Example: N=5, values: 0, 2, 4, 6, 8, n=2 (with replacement).
25 possible samples.

Prob. of each samples
mean = 1/25 = 0.04

?	?	?	?	?
?	?	?	?	?
?	?	?	?	?
?	?	?	?	?
?	?	?	?	?

$\bar{x} = ?$, occurs xxx times

$\bar{x} = ?$, occurs xxx times

$\bar{x} = ?$, occurs xxxxx times

$\bar{x} = ?$, occurs xxxx times

$\bar{x} = ?$, occurs xxxx times

$\bar{x} = ?$, occurs xxxx times

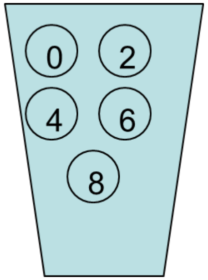
$\bar{x} = ?$, occurs xxxxx times

$\bar{x} = ?$, occurs xxx times

$\bar{x} = ?$, occurs xxx times

7: Sample Variability

7.2 The Sampling Distribution of Sample Means



Example: N=5, values: 0, 2, 4, 6, 8, n=2 (with replacement).
25 possible samples.

Prob. of each samples
mean = 1/25 = 0.04

0	1	2	3	4
1	2	3	4	5
2	3	4	5	6
3	4	5	6	7
4	5	6	7	8

$\bar{x} = 0,$ occurs one time

$\bar{x} = 1,$ occurs two times

$\bar{x} = 2,$ occurs three times

$\bar{x} = 3,$ occurs four times

$\bar{x} = 4,$ occurs five times

$\bar{x} = 5,$ occurs four times

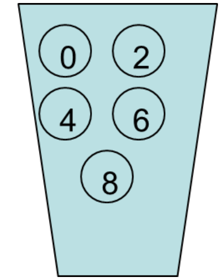
$\bar{x} = 6,$ occurs three times

$\bar{x} = 7,$ occurs two times

$\bar{x} = 8,$ occurs one time

7: Sample Variability

7.2 The Sampling Distribution of Sample Means



Example: N=5, values: 0, 2, 4, 6, 8, n=2 (with replacement).
25 possible samples.

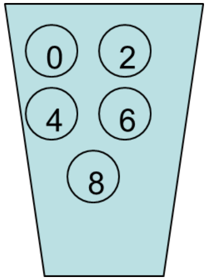
Prob. of each samples mean = 1/25 = 0.04

?	?	?	?	?
?	?	?	?	?
?	?	?	?	?
?	?	?	?	?
?	?	?	?	?

$$P(\bar{x} = ?) =$$
$$P(\bar{x} = ?) =$$
$$P(\bar{x} = ?) =$$
$$P(\bar{x} = ?) =$$
$$P(\bar{x} = ?) =$$
$$P(\bar{x} = ?) =$$
$$P(\bar{x} = ?) =$$
$$P(\bar{x} = ?) =$$
$$P(\bar{x} = ?) =$$

7: Sample Variability

7.2 The Sampling Distribution of Sample Means



Example: N=5, values: 0, 2, 4, 6, 8, n=2 (with replacement).
25 possible samples.

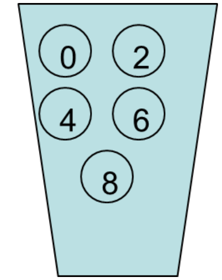
Prob. of each samples mean = 1/25 = 0.04

0	1	2	3	4
1	2	3	4	5
2	3	4	5	6
3	4	5	6	7
4	5	6	7	8

$$P(\bar{x}=0)=1/25$$
$$P(\bar{x}=1)=2/25$$
$$P(\bar{x}=2)=3/25$$
$$P(\bar{x}=3)=4/25$$
$$P(\bar{x}=4)=5/25$$
$$P(\bar{x}=5)=4/25$$
$$P(\bar{x}=6)=3/25$$
$$P(\bar{x}=7)=2/25$$
$$P(\bar{x}=8)=1/25$$

7: Sample Variability

7.2 The Sampling Distribution of Sample Means



Example: N=5, values: 0, 2, 4, 6, 8, n=2 (with replacement).

$$P(\bar{x}=0)=1/25$$
$$P(\bar{x}=1)=2/25$$
$$P(\bar{x}=2)=3/25$$
$$P(\bar{x}=3)=4/25$$
$$P(\bar{x}=4)=5/25$$
$$P(\bar{x}=5)=4/25$$
$$P(\bar{x}=6)=3/25$$
$$P(\bar{x}=7)=2/25$$
$$P(\bar{x}=8)=1/25$$

Represent this probability distribution with a histogram.

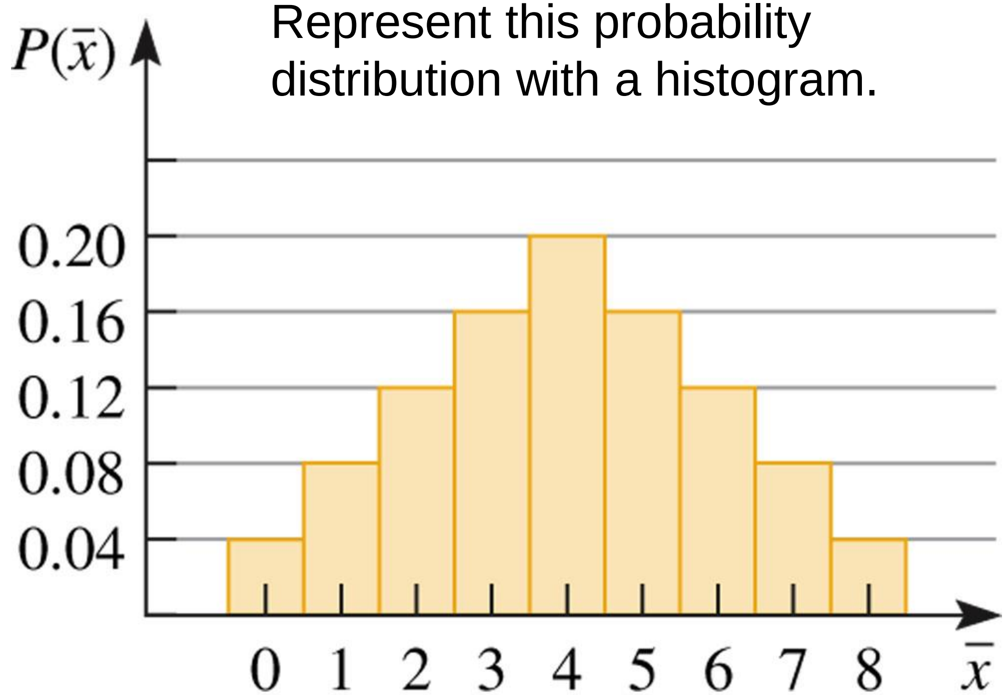
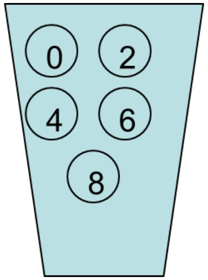


Figure from Johnson & Kuby, 2012.

7: Sample Variability

7.2 The Sampling Distribution of Sample Means



Don't forget that the two values that we draw are random.

That is, we may know the sample space of possible outcomes

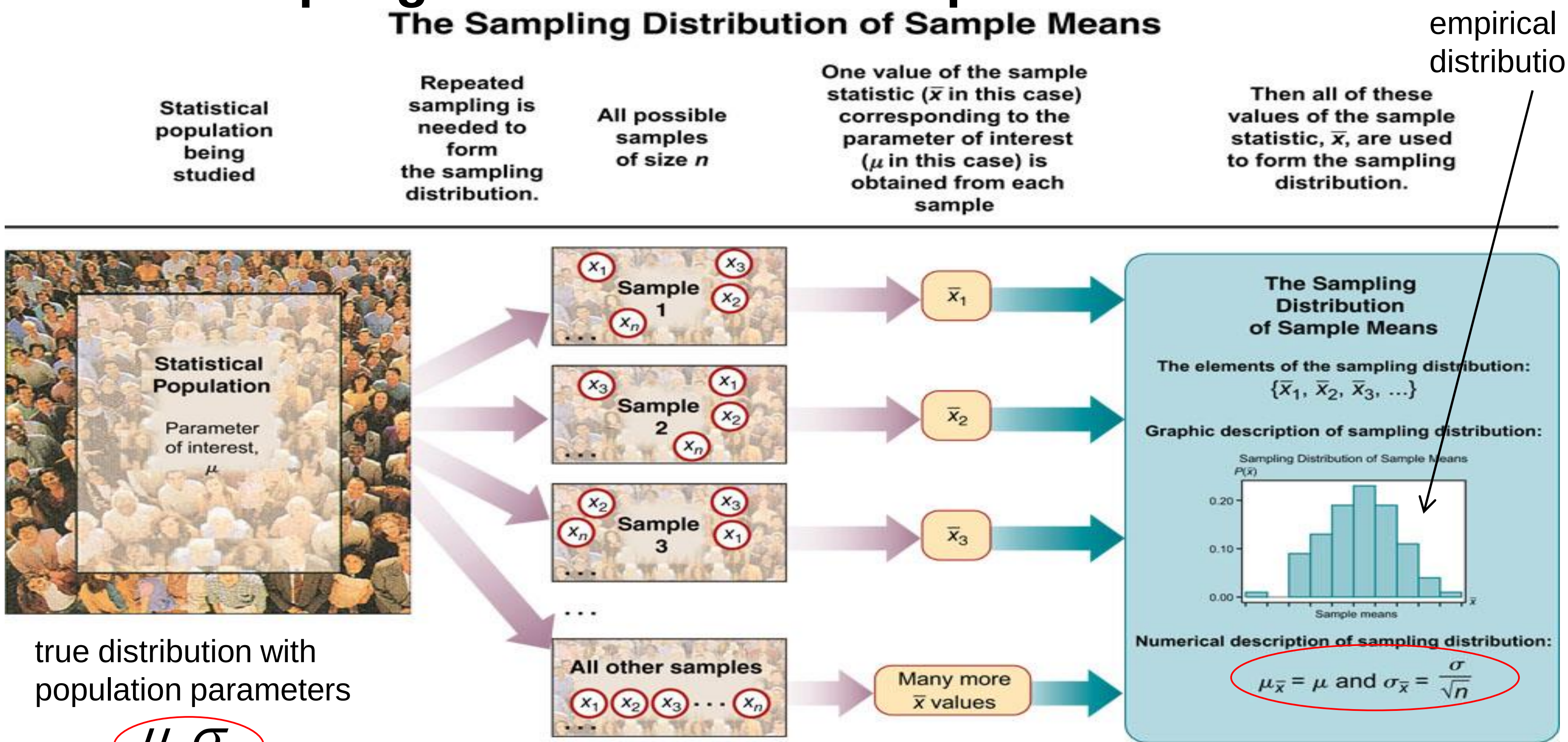
but we do not know exactly which ones we will get!

Random Sample: A sample obtained in such a way that each possible sample of fixed size n has an equal probability of being selected.

7: Sample Variability

7.2 The Sampling Distribution of Sample Means

empirical
distribution



true distribution with
population parameters

μ, σ

Figure from Johnson & Kuby, 2012.

As the number of samples increases the empirical dist. turns into theoretical dist.

7: Sample Variability

7.2 The Sampling Distribution of Sample Means

Sample distribution of sample means (SDSM): If all possible random samples, each of size n, are taken from any population with mean μ and standard deviation σ, then the sampling distribution of sample means will have the following:

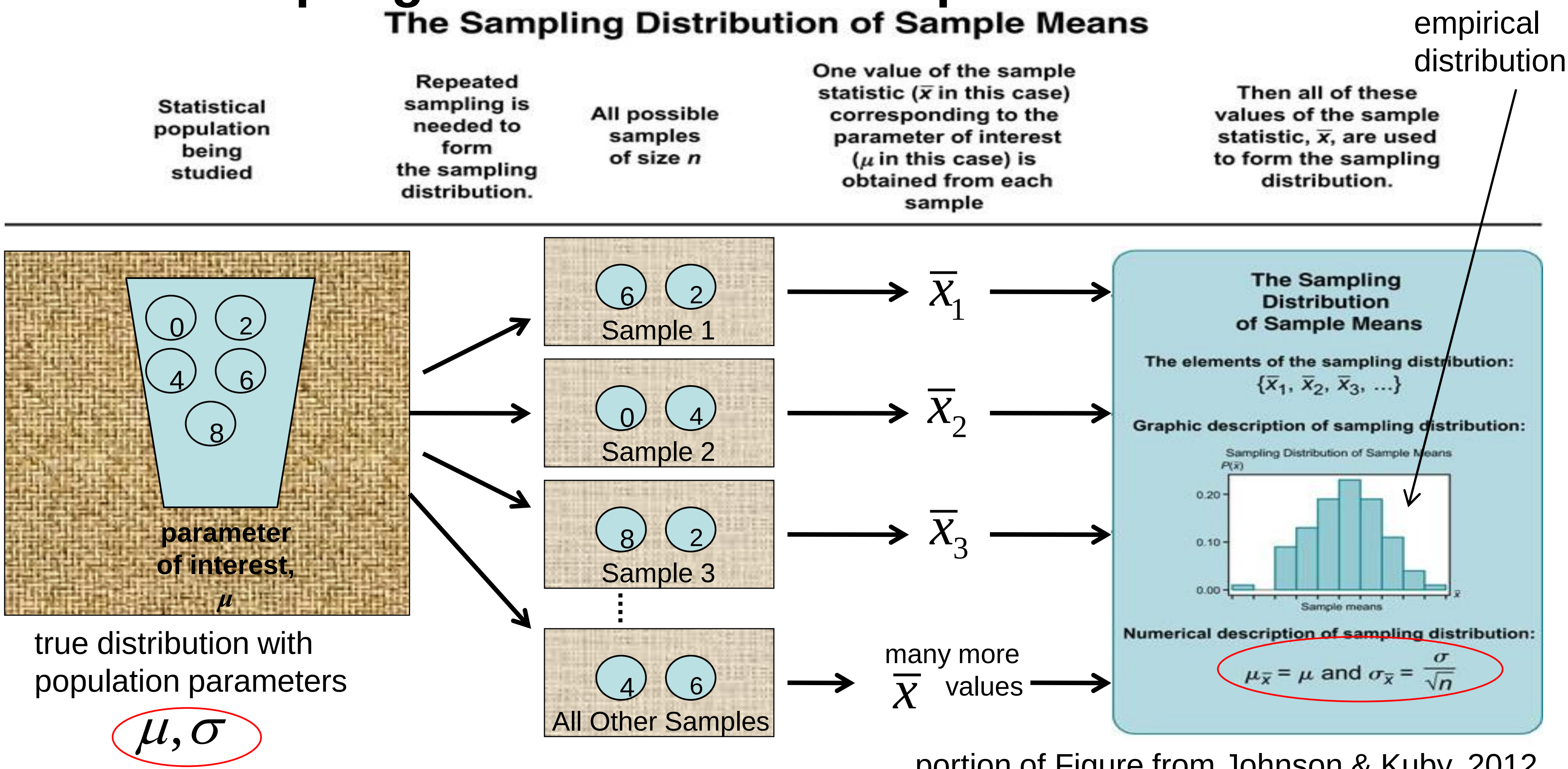
1. A mean $\mu_{\bar{x}}$ equal to μ
2. A standard deviation $\sigma_{\bar{x}}$ equal to $\dfrac{\sigma}{\sqrt{n}}$

Furthermore, if the sampled population has a normal distribution, then the sampling distribution of $\bar{x}$ will also be normal for all samples of all sizes.

Discuss Later: What if the sampled population does not have a normal distribution?

7: Sample Variability

7.2 The Sampling Distribution of Sample Means

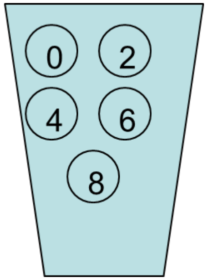


portion of Figure from Johnson & Kuby, 2012.

As the number of samples increases the empirical dist. turns into theoretical dist.

7: Sample Variability

7.2 The Sampling Distribution of Sample Means

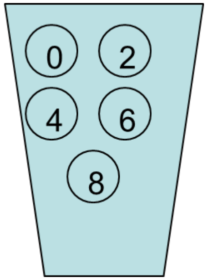


Example: N=5, values: 0, 2, 4, 6, 8, n=2 (with replacement).
Instead of drawing two values with replacement and computing the sample mean, we can think of this as drawing one of the sample means with replacement.

The probability for each sample mean is →

7: Sample Variability

7.2 The Sampling Distribution of Sample Means



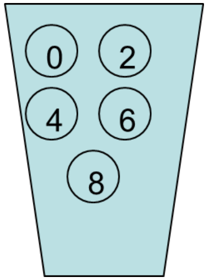
Example: N=5, values: 0, 2, 4, 6, 8, n=2 (with replacement).
Instead of drawing two values with replacement and computing the sample mean, we can think of this as drawing one of the sample means with replacement.

The probability for each sample mean is →

$$P(\bar{x} = 0) = 1/25$$
$$P(\bar{x} = 1) = 2/25$$
$$P(\bar{x} = 2) = 3/25$$
$$P(\bar{x} = 3) = 4/25$$
$$P(\bar{x} = 4) = 5/25$$
$$P(\bar{x} = 5) = 4/25$$
$$P(\bar{x} = 6) = 3/25$$
$$P(\bar{x} = 7) = 2/25$$
$$P(\bar{x} = 8) = 1/25$$

7: Sample Variability

7.2 The Sampling Distribution of Sample Means



Example: N=5, values: 0, 2, 4, 6, 8, n=2 (with replacement).

$$\mu_{\bar{x}} = \sum \bar{x} P(\bar{x})$$

$$\mu_{\bar{x}} =$$

$$P(\bar{x} = ?) =$$

$$P(\bar{x} = ?) =$$

$$P(\bar{x} = ?) =$$

$$P(\bar{x} = ?) =$$

$$P(\bar{x} = ?) =$$

$$P(\bar{x} = ?) =$$

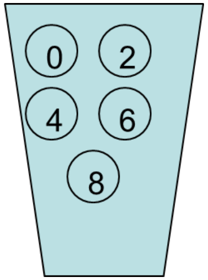
$$P(\bar{x} = ?) =$$

$$P(\bar{x} = ?) =$$

$$P(\bar{x} = ?) =$$

7: Sample Variability

7.2 The Sampling Distribution of Sample Means



Example: N=5, values: 0, 2, 4, 6, 8, n=2 (with replacement).

$$\mu_{\bar{x}} = \sum \bar{x} P(\bar{x})$$

$$\mu_{\bar{x}} = 0(1/25) + 1(2/25)$$
$$+ 2(3/25) + 3(4/25)$$
$$+ 4(5/25) + 5(4/25)$$
$$+ 6(3/25) + 7(2/25)$$
$$+ 8(1/25)$$

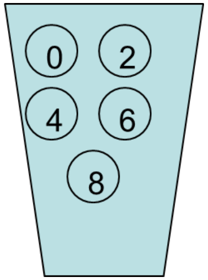
$$\mu_{\bar{x}} = 4$$ ← Same as SDSM formula!

↘ $$\mu_{\bar{x}} = \mu = 4$$

$$P(\bar{x} = 0) = 1/25$$
$$P(\bar{x} = 1) = 2/25$$
$$P(\bar{x} = 2) = 3/25$$
$$P(\bar{x} = 3) = 4/25$$
$$P(\bar{x} = 4) = 5/25$$
$$P(\bar{x} = 5) = 4/25$$
$$P(\bar{x} = 6) = 3/25$$
$$P(\bar{x} = 7) = 2/25$$
$$P(\bar{x} = 8) = 1/25$$

7: Sample Variability

7.2 The Sampling Distribution of Sample Means



Example: N=5, values: 0, 2, 4, 6, 8, n=2 (with replacement).

$$\sigma_{\bar{x}}^2 = \sum (\bar{x} - \mu)^2 P(\bar{x})$$

$$\sigma_{\bar{x}}^2 =$$

$$P(\bar{x} = ?) =$$

$$P(\bar{x} = ?) =$$

$$P(\bar{x} = ?) =$$

$$P(\bar{x} = ?) =$$

$$P(\bar{x} = ?) =$$

$$P(\bar{x} = ?) =$$

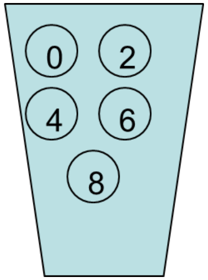
$$P(\bar{x} = ?) =$$

$$P(\bar{x} = ?) =$$

$$P(\bar{x} = ?) =$$

7: Sample Variability

7.2 The Sampling Distribution of Sample Means



Example: N=5, values: 0, 2, 4, 6, 8, n=2 (with replacement).

$$\sigma_{\bar{x}}^2 = \sum (\bar{x}-\mu)^2 P(\bar{x})$$

$$\begin{aligned}\sigma_{\bar{x}}^2 = &(0-4)^2(1/25)+(1-4)^2(2/25)\\ &+(2-4)^2(3/25)+(3-4)^2(4/25)\\ &+(4-4)^2(5/25)+(5-4)^2(4/25)\\ &+(6-4)^2(3/25)+(7-4)^2(2/25)\\ &+(8-4)^2(1/25)\end{aligned}$$

$$\sigma_{\bar{x}}^2 = 4$$

$$\sigma_{\bar{x}} = 2$$

Same as SDSM formula!

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{2\sqrt{2}}{\sqrt{2}} = 2$$

$$P(\bar{x}=0)=1/25$$
$$P(\bar{x}=1)=2/25$$
$$P(\bar{x}=2)=3/25$$
$$P(\bar{x}=3)=4/25$$
$$P(\bar{x}=4)=5/25$$
$$P(\bar{x}=5)=4/25$$
$$P(\bar{x}=6)=3/25$$
$$P(\bar{x}=7)=2/25$$
$$P(\bar{x}=8)=1/25$$

7: Sample Variability

7.2 The Sampling Distribution of Sample Means

Sample distribution of sample means (SDSM): If all possible random samples, each of size n, are taken from any population with mean μ and standard deviation σ, then the sampling distribution of sample means will have the following:

1. A mean $\mu_{\bar{x}}$ equal to μ
2. A standard deviation $\sigma_{\bar{x}}$ equal to $\dfrac{\sigma}{\sqrt{n}}$

Furthermore, if the sampled population has a normal distribution, then the sampling distribution of $\bar{x}$ will also be normal for all samples of all sizes.

Discuss Later: What if the sampled population does not have a normal distribution?

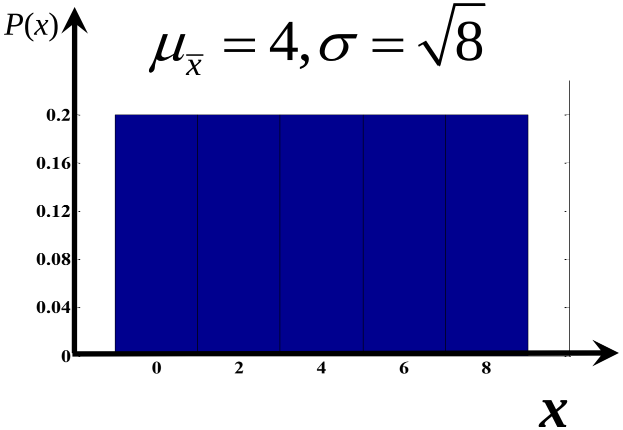
1. A mean $\mu_{\bar{x}}$ equal to μ
2. A standard deviation $\sigma_{\bar{x}}$ equal to $\dfrac{\sigma}{\sqrt{n}}$

7: Sample Variability

7.2 The Sampling Distribution of Sample Means

$\bar{x}$ from n=1 distribution

$\mu_{\bar{x}} = 4, \sigma = \sqrt{8}$



$\bar{x}$ from n=2 distribution

$\mu_{\bar{x}} = 4, \sigma_{\bar{x}} = \sqrt{8}/\sqrt{2}$

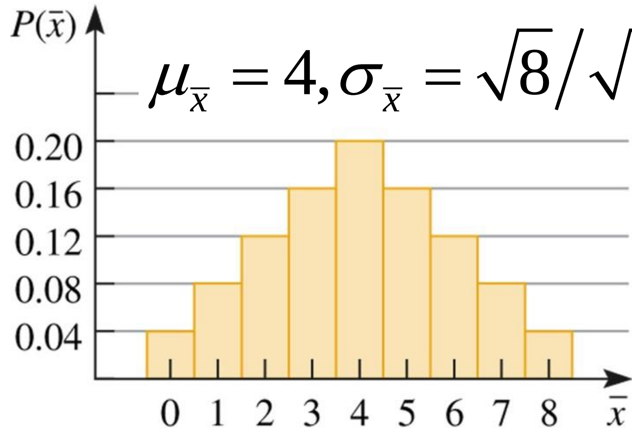
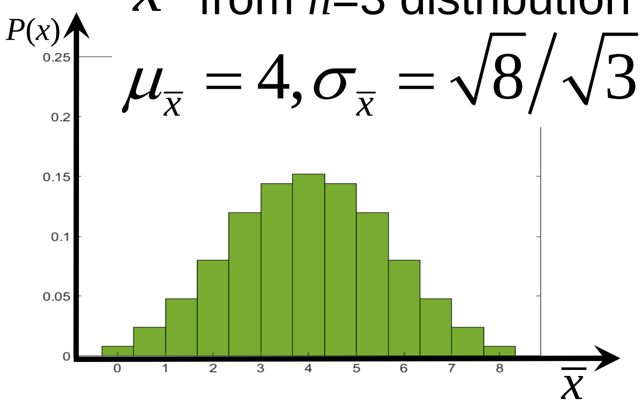


Figure from Johnson & Kuby, 2012.

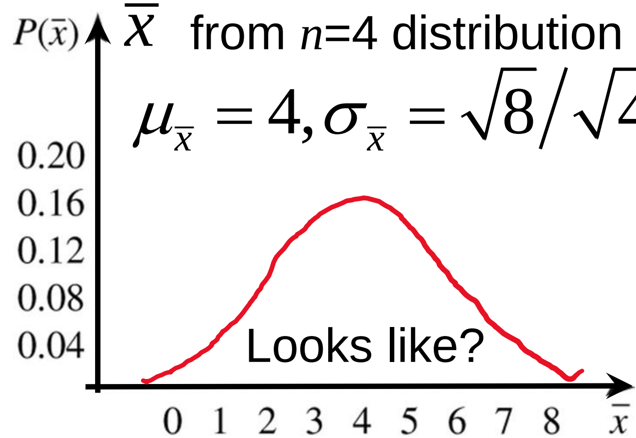
$\bar{x}$ from n=3 distribution

$\mu_{\bar{x}} = 4, \sigma_{\bar{x}} = \sqrt{8}/\sqrt{3}$



$\bar{x}$ from n=4 distribution

$\mu_{\bar{x}} = 4, \sigma_{\bar{x}} = \sqrt{8}/\sqrt{4}$



n large?

$\mu_{\bar{x}} = 4$

$\sigma_{\bar{x}} = \sqrt{8}/\sqrt{n}$

Looks like?

See end of slides for details

7: Sample Variability

7.2 The Sampling Distribution of Sample Means

We have a couple of definitions.

Standard error of the mean ($\sigma_{\bar{x}}$): The standard deviation of the sampling distribution of sample means.

Central Limit Theorem (CLT): The sampling distribution of sample means will more closely resemble the normal distribution as the sample size increases.

The *CLT* is **extremely** important in Statistics!

7: Sample Variability

7.2 The Sampling Distribution of Sample Means

The Central Limit Theorem: Assume that we have a population (arbitrary distribution) with mean μ and standard deviation σ.

If we take random samples of size n (with replacement), then for "large" n, the distribution of the sample means, the $\bar{x}$'s, is approximately normally distributed with

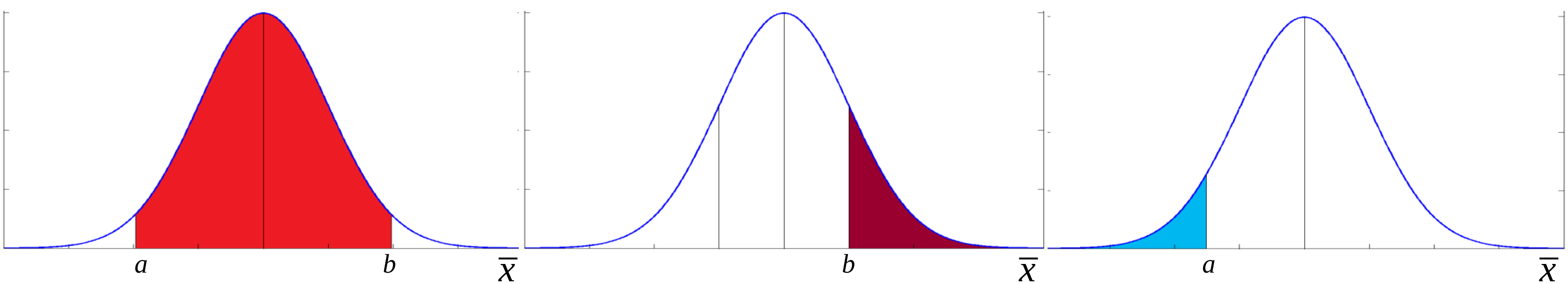
$$\mu_{\bar{x}} = \mu, \qquad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

where in general $n \geq 30$ is sufficiently "large," but can be as small as15 or as big as 50 depending upon the shape of distribution!

7: Sample Variability

7.3 Application of the Sampling Distribution of Sample Means

Now that we believe that the mean $\bar{x}$ from a sample of n=15 is normally distributed with mean $\mu_{\bar{x}} = \mu$ and standard deviation $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$, we can find probabilities.

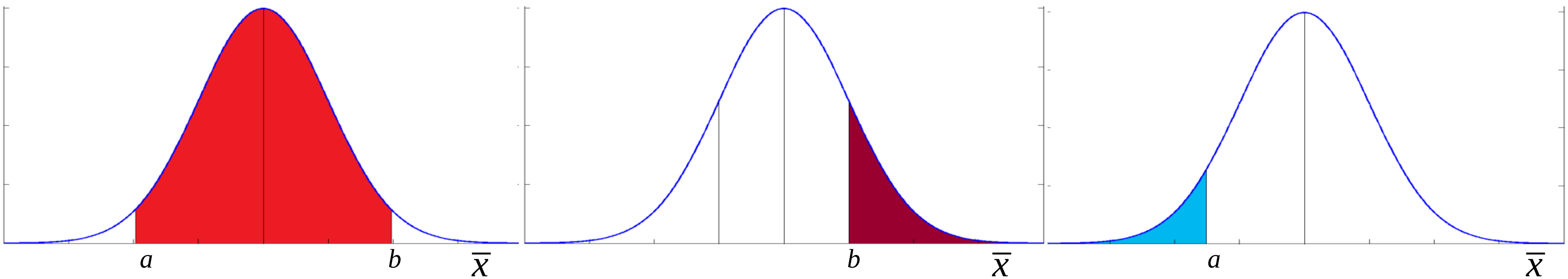


$P(a < \bar{x} < b)$ $P(b < \bar{x})$ $P(\bar{x} < a)$

7: Sample Variability

7.3 Application of the Sampling Distribution of Sample Means

To find these probabilities, we first convert to z scores



$$P(a<\bar{x}<b)$$ same area $$P(c<z<d)$$

$$P(b<\bar{x})$$ same area $$P(d<z)$$

$$P(\bar{x}<a)$$ same area $$P(z<c)$$

$$z=\frac{\bar{x}-\mu_{\bar{x}}}{\sigma_{\bar{x}}},\quad c=\frac{a-\mu_{\bar{x}}}{\sigma_{\bar{x}}},\quad d=\frac{b-\mu_{\bar{x}}}{\sigma_{\bar{x}}}$$ and use the table in book.

7: Sample Variability

7.3 Application of the Sampling Distribution of Sample Means

Example:

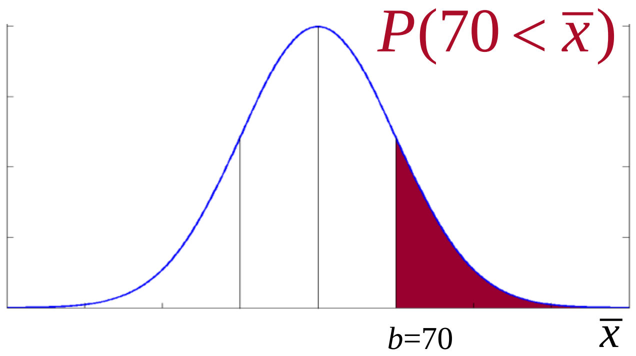
What is probability that sample mean $\bar{x}$ from a random sample of n=15 heights is greater than 70" when $\mu = 67.0$ and $\sigma = 4.3$?

7: Sample Variability

7.3 Application of the Sampling Distribution of Sample Means

Example:

What is probability that sample mean $\bar{x}$ from a random sample of n=15 heights is greater than 70” when $\mu = 67.0$ and $\sigma = 4.3$?

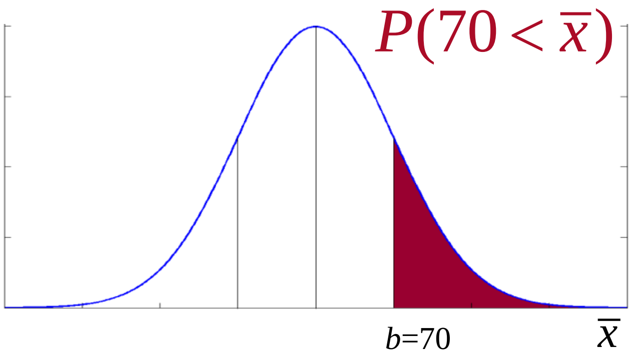


7: Sample Variability

7.3 Application of the Sampling Distribution of Sample Means

Example:

What is probability that sample mean $\bar{x}$ from a random sample of n=15 heights is greater than 70" when $\mu = 67.0$ and $\sigma = 4.3$?



we first convert to z scores

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}}$$

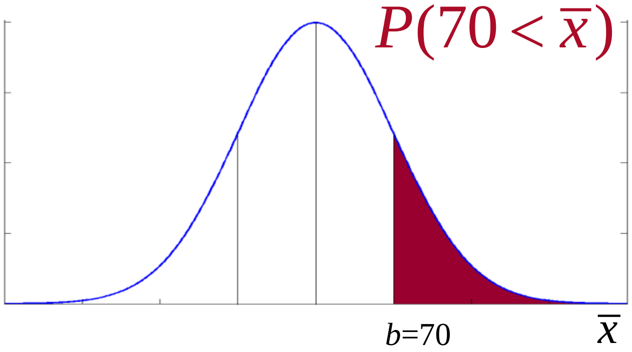
$$d = \frac{b - \mu_{\bar{x}}}{\sigma_{\bar{x}}}$$

7: Sample Variability

7.3 Application of the Sampling Distribution of Sample Means

Example:

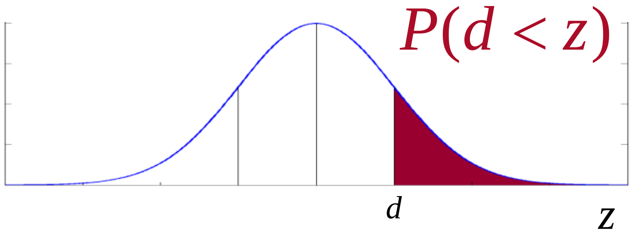
What is probability that sample mean $\bar{x}$ from a random sample of n=15 heights is greater than 70" when $\mu = 67.0$ and $\sigma = 4.3$?

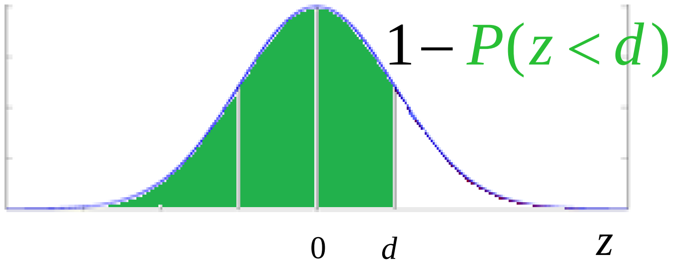


we first convert to z scores

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}}$$

$$d = \frac{b - \mu_{\bar{x}}}{\sigma_{\bar{x}}}$$



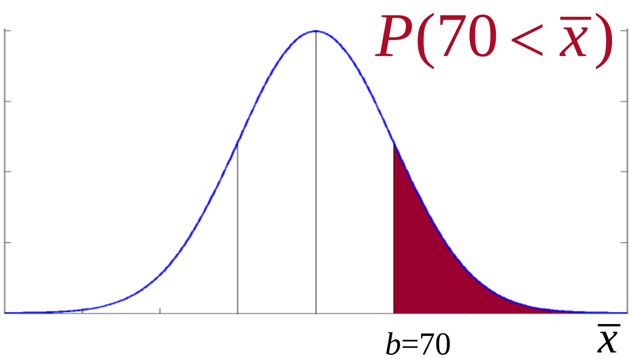


7: Sample Variability

7.3 Application of the Sampling Distribution of Sample Means

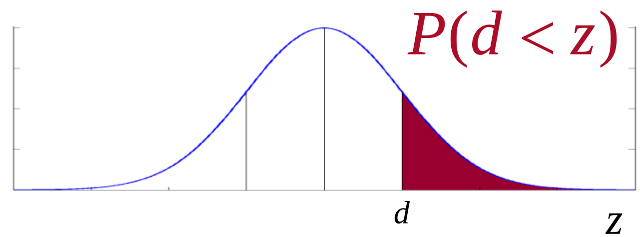
Example:

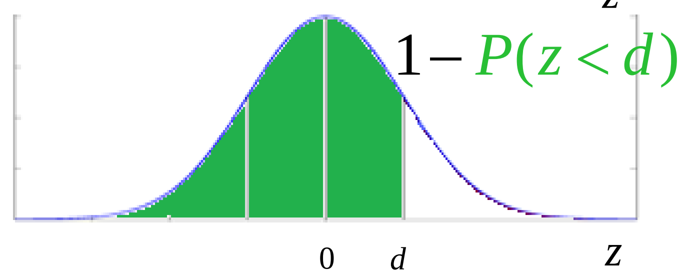
What is probability that sample mean $\bar{x}$ from a random sample of n=15 heights is greater than 70” when $\mu = 67.0$ and $\sigma = 4.3$?



we first convert to z scores

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}}$$





where $d = \dfrac{b - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \dfrac{70 - 67.0}{4.3 / \sqrt{15}} = 2.70$, then use the table in book.

$$1 - P(z < 2.70) = 1 - .9965 = 0.0035$$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974

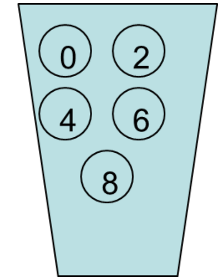
7: Sample Variability

Questions?

Homework: Read Chapter 7.1-7.3
WebAssign
Chapter 7 # 6, 21, 23, 29, 33, 35

7: Sample Variability

7.2 The Sampling Distribution of Sample Means



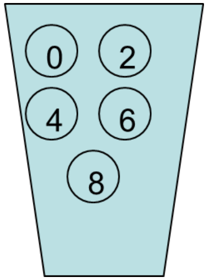
Example: N=5, values: 0, 2, 4, 6, 8, n=3 (with replacement).
125 possible samples.

(0,0,0)	(2,0,0)	(4,0,0)	(6,0,0)	(8,0,0)	(0,0,2)	(2,0,2)	(4,0,2)	(6,0,2)	(8,0,2)	(0,0,4)	(2,0,4)	(4,0,4)	(6,0,4)	(8,0,4)
(0,2,0)	(2,2,0)	(4,2,0)	(6,2,0)	(8,2,0)	(0,2,2)	(2,2,2)	(4,2,2)	(6,2,2)	(8,2,2)	(0,2,4)	(2,2,4)	(4,2,4)	(6,2,4)	(8,2,4)
(0,4,0)	(2,4,0)	(4,4,0)	(6,4,0)	(8,4,0)	(0,4,2)	(2,4,2)	(4,4,2)	(6,4,2)	(8,4,2)	(0,4,4)	(2,4,4)	(4,4,4)	(6,4,4)	(8,4,4)
(0,6,0)	(2,6,0)	(4,6,0)	(6,6,0)	(8,6,0)	(0,6,2)	(2,6,2)	(4,6,2)	(6,6,2)	(8,6,2)	(0,6,4)	(2,6,4)	(4,6,4)	(6,6,4)	(8,6,4)
(0,8,0)	(2,8,0)	(4,8,0)	(6,8,0)	(8,8,0)	(0,8,2)	(2,8,2)	(4,8,2)	(6,8,2)	(8,8,2)	(0,8,4)	(2,8,4)	(4,8,4)	(6,8,4)	(8,8,4)

(0,0,6)	(2,0,6)	(4,0,6)	(6,0,6)	(8,0,6)	(0,0,8)	(2,0,8)	(4,0,8)	(6,0,8)	(8,0,8)
(0,2,6)	(2,2,6)	(4,2,6)	(6,2,6)	(8,2,6)	(0,2,8)	(2,2,8)	(4,2,8)	(6,2,8)	(8,2,8)
(0,4,6)	(2,4,6)	(4,4,6)	(6,4,6)	(8,4,6)	(0,4,8)	(2,4,8)	(4,4,8)	(6,4,8)	(8,4,8)
(0,6,6)	(2,6,6)	(4,6,6)	(6,6,6)	(8,6,6)	(0,6,8)	(2,6,8)	(4,6,8)	(6,6,8)	(8,6,8)
(0,8,6)	(2,8,6)	(4,8,6)	(6,8,6)	(8,8,6)	(0,8,8)	(2,8,8)	(4,8,8)	(6,8,8)	(8,8,8)

7: Sample Variability

7.2 The Sampling Distribution of Sample Means



Example: N=5, values: 0, 2, 4, 6, 8, n=3 (with replacement).
125 possible samples.

0	2 / 3	4 / 3	6 / 3	8 / 3
2 / 3	4 / 3	6 / 3	8 / 3	10 / 2
4 / 3	6 / 3	8 / 3	10 / 2	12 / .2
6 / 3	8 / 3	10 / 2	12 / 2	14 / 2
8 / 3	10 / 2	12 / 2	14 / 2	16 / 2

2 / 3	4 / 3	6 / 3	8 / 3	10 / 3
4 / 3	6 / 3	8 / 3	10 / 3	12 / 3
6 / 3	8 / 3	10 / 3	12 / 3	14 / 3
8 / 3	10 / 3	12 / 3	14 / 3	16 / 3
10 / 3	12 / 3	14 / 3	16 / 3	18 / 3

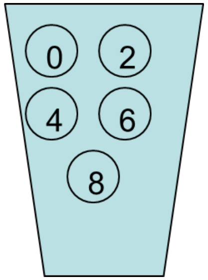
4 / 3	6 / 3	8 / 3	10 / 3	12 / 3
6 / 3	8 / 3	10 / 3	12 / 3	14 / 3
8 / 3	10 / 3	12 / 3	14 / 3	16 / 3
10 / 3	12 / 3	14 / 3	16 / 3	18 / 3
12 / 3	14 / 3	16 / 3	18 / 3	20 / 3

6 / 3	8 / 3	10 / 3	12 / 3	14 / 3
8 / 3	10 / 3	12 / 3	14 / 3	16 / 3
10 / 3	12 / 3	14 / 3	16 / 3	18 / 3
12 / 3	14 / 3	16 / 3	18 / 3	20 / 3
14 / 3	16 / 3	18 / 3	20 / 3	22 / 3

8 / 3	10 / 3	12 / 3	14 / 3	16 / 3
10 / 3	12 / 3	14 / 3	16 / 3	18 / 3
12 / 3	14 / 3	16 / 3	18 / 3	20 / 3
14 / 3	16 / 3	18 / 3	20 / 3	22 / 3
16 / 3	18 / 3	20 / 3	22 / 3	24 / 3

7: Sample Variability

7.2 The Sampling Distribution of Sample Means

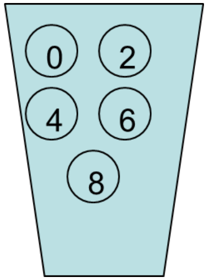


Example: N=5, values: 0, 2, 4, 6, 8, n=3 (with replacement).
125 possible samples.

$$P(\bar{x} = 0/3) = 1/125$$
$$P(\bar{x} = 2/3) = 3/125$$
$$P(\bar{x} = 4/3) = 3/125$$
$$P(\bar{x} = 6/3) = 10/125$$
$$P(\bar{x} = 8/3) = 15/125$$
$$P(\bar{x} = 10/3) = 18/125$$
$$P(\bar{x} = 12/3) = 19/125$$
$$P(\bar{x} = 14/3) = 18/125$$
$$P(\bar{x} = 16/3) = 15/125$$
$$P(\bar{x} = 18/3) = 10/125$$
$$P(\bar{x} = 20/3) = 6/125$$
$$P(\bar{x} = 22/3) = 3/125$$
$$P(\bar{x} = 24/3) = 1/125$$

7: Sample Variability

7.2 The Sampling Distribution of Sample Means



Example: N=5, values: 0, 2, 4, 6, 8, n=3 (with replacement).

$P(\bar{x} = 0/3) = 1/125$

$P(\bar{x} = 2/3) = 3/125$

$P(\bar{x} = 4/3) = 3/125$

$P(\bar{x} = 6/3) = 10/125$

$P(\bar{x} = 8/3) = 15/125$

$P(\bar{x} = 10/3) = 18/125$

$P(\bar{x} = 12/3) = 19/125$

$P(\bar{x} = 14/3) = 18/125$

$P(\bar{x} = 16/3) = 15/125$

$P(\bar{x} = 18/3) = 10/125$

$P(\bar{x} = 20/3) = 6/125$

$P(\bar{x} = 22/3) = 3/125$

$P(\bar{x} = 24/3) = 1/125$

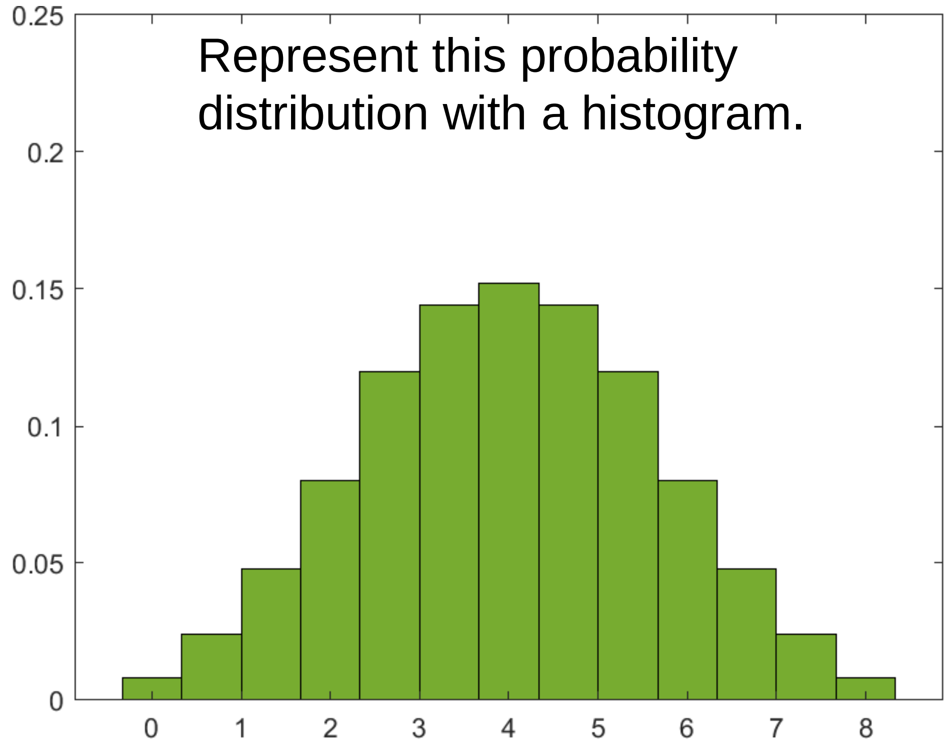


Figure from Johnson & Kuby, 2012.