Unit II: Inferential Statistics – I

WHAT *IS A SAMPLING DISTRIBUTION?*

Sampling distribution: a distribution of statistics obtained by selecting all the possible samples of a specific size from a population.

Distribution of Sample Means: the collection of sample means for **all** the possible random samples of a **particular** size (n) that can be obtained from a population

THE CENTRAL LIMIT THEOREM

The central limit theorem states that
(1) as the size of the sample (n) increases, the shape of the sampling distribution of means approximates the shape of the normal distribution;
(2) the mean of the sampling distribution of means will equal the population mean (μ); and
(3) the standard deviation will equal $\sigma / \sqrt{n}$

Shape. Even if the shape of the population distribution from which a sample is drawn is not normal, if the sample size is large (i.e., n = 60 or more), the sampling distribution itself will be normal in shape.

Mean: The mean of the distribution of sample means is the mean of the population. Sample size does not affect the center of the distribution.

Standard Deviation. The standard deviation of the sampling distribution of means is called the **standard error of the mean (σM), or simply standard error.** The standard error (σM) represents the amount that a sample mean (M) is expected to vary from the population mean (μ).

$$\sigma_M = \frac{\sigma}{\sqrt{n}}.$$

Give: $\sigma = 16$ $\quad$ $n = 25$

$$\sigma_M = \frac{16}{\sqrt{25}} = 3.2$$

$\sigma = 16$ $\quad$ $n = 100$

$$\sigma_M = \frac{16}{\sqrt{100}} = 1.6$$

Thus, larger samples generally produce less sampling error.

PROBABILITIES, PROPORTIONS, AND PERCENTAGES OF SAMPLE MEANS

The sampling distribution of means can be used to determine the probabilities, proportions, and percentages associated with particular sample means.

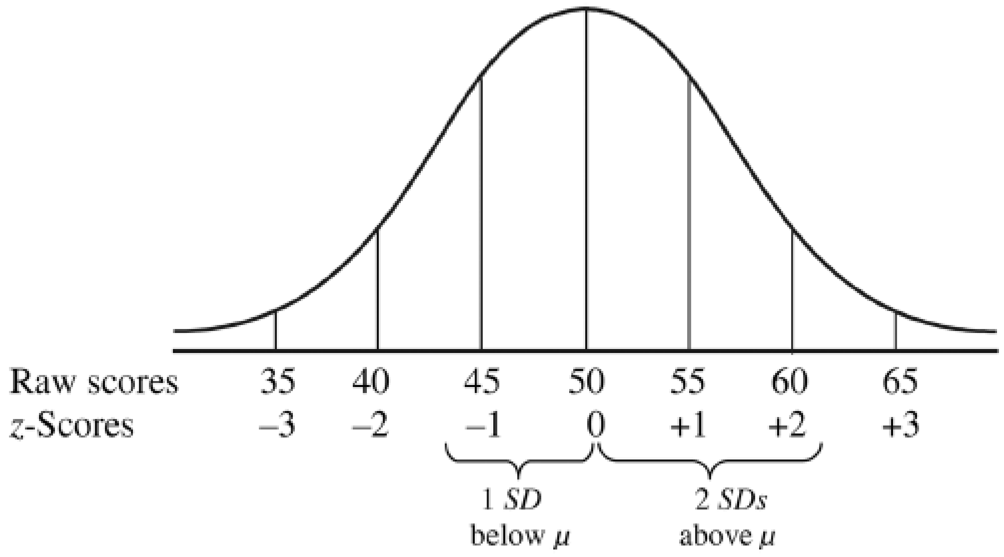
Use the normal distribution curve to determine these figures for particular raw scores.

$$z = \frac{M - \mu}{\sigma_M} \quad \text{and} \quad M = \mu + (z)(\sigma_M)$$

PROBABILITIES, PROPORTIONS, AND PERCENTAGES OF SAMPLE MEANS

z-score, describes how far a particular raw score deviates from the mean in standard deviation units.

The relationship between raw scores and *z-scores is illustrated below*

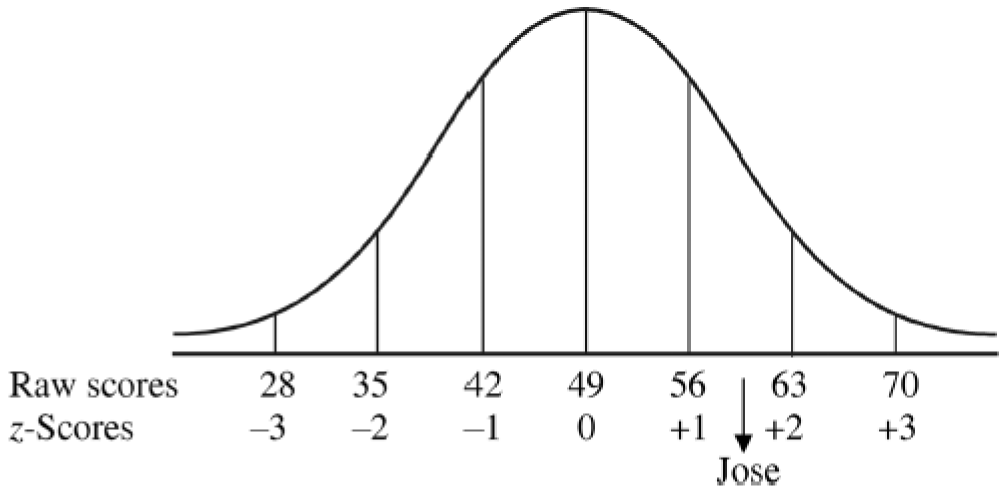


$$z = \frac{X - \mu}{\sigma}$$

Jose scored 60 on his history test. The class mean was 49, and the standard deviation was 7. What is Jose's equivalent *z*-score?

$$z = \frac{60 - 49}{7} = +1.57$$

Thus, a raw score of 60 converts to a *z*-score of +1.57, meaning that Jose scored 1.57 standard deviations above the mean. This situation is illustrated in the following graph:



THE *Z-TABLE*

Column A – indicates a particular *z-score value.*

Column B – indicates the proportion in the body of the curve (i.e., the larger portion).

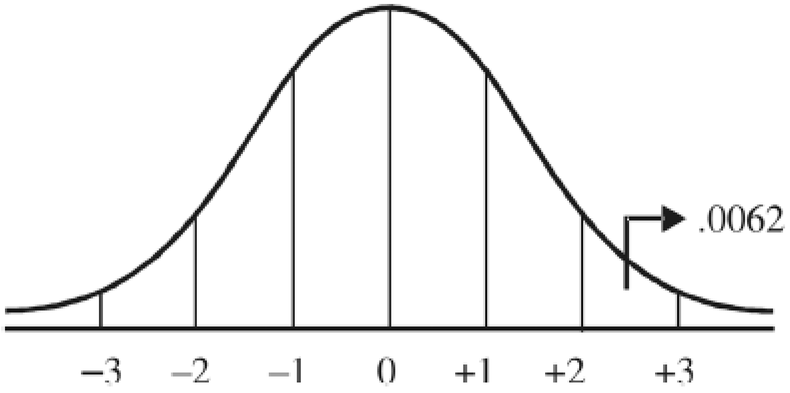
Column C – indicates the proportion in the tail (i.e., the smaller portion).

Column D – indicates the proportion between the mean and the z-score value in Column A. (Remember that the mean of a z-scale is 0 and it is located in the center of the distribution.)

To find *p* values in the *z*-table that are:

- above a positive number, look in Column C (tail).
- below a positive number, look in Column B (body).
- number, look in Column B (body).
- below a negative number, look in Column C (tail).

What proportion of the normal distribution is associated with *z-scores greater than +2.50? The proper notation for this question is: p(z > + 2.50) = ?*



$$p(z > +2.50) = .0062$$

Other Examples Are as Follows

$p(z < +2.00) = .9772$ (look in the body, Column B)
$p(z < -1.50) = .0668$ (look in the tail, Column C)
$p(z > -1.23) = .8907$ (look in the body, Column B)

Column A	*Column B*	*Column C*	*Column D*	*Column A*	*Column B*	*Column C*	*Column D*
z-score	*Proportion in body (larger part)*	*Proportion in tail (smaller part)*	*Proportion between mean and z*	*z-score*	*Proportion in body (larger part)*	*Proportion in tail (smaller part)*	*Proportion between mean and z*
1.94	.9738	.0262	.4738	**2.45**	.9929	.0071	.4929
1.95	.9744	.0256	.4744	**2.46**	.9931	.0069	.4931
1.96	.9750	.0250	.4750	**2.47**	.9932	.0068	.4932
1.97	.9756	.0244	.4756	**2.48**	.9934	.0066	.4934
1.98	.9761	.0239	.4761	**2.49**	.9936	.0064	.4936
1.99	.9767	.0233	.4767	**2.50**	.9938	.0062	.4938
2.00	.9772	.0228	.4772	**2.51**	.9940	.0060	.4940
2.01	.9778	.0222	.4778	**2.52**	.9941	.0059	.4941
2.02	.9783	.0217	.4783	**2.53**	.9943	.0057	.4943
2.03	.9788	.0212	.4788	**2.54**	.9945	.0055	.4945
2.04	.9793	.0207	.4793	**2.55**	.9946	.0054	.4946
2.05	.9798	.0202	.4798	**2.56**	.9948	.0052	.4948
2.06	.9803	.0197	.4803	**2.57**	.9949	.0051	.4949
2.07	.9808	.0192	.4808	**2.58**	.9951	.0049	.4951
2.08	.9812	.0188	.4812	**2.59**	.9952	.0048	.4952
2.09	.9817	.0183	.4817	**2.60**	.9953	.0047	.4953

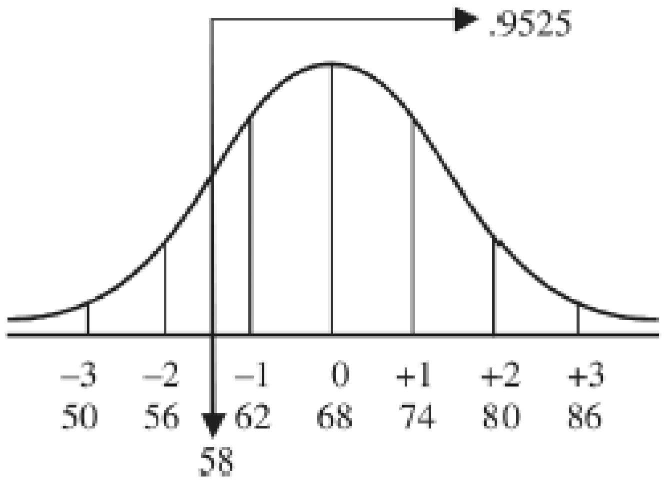
1. What is the *probability* of obtaining a score greater than 58? Because we are looking for the probability associated with a raw score (X) greater than 58, the correct notation for this problem would be:

$p(X > 58) =$

The raw score will first have to be converted to a *z*-score so that you can refer to the *z*-table.

$$z = \frac{X - \mu}{\sigma} = \frac{58 - 68}{6} = -1.67$$

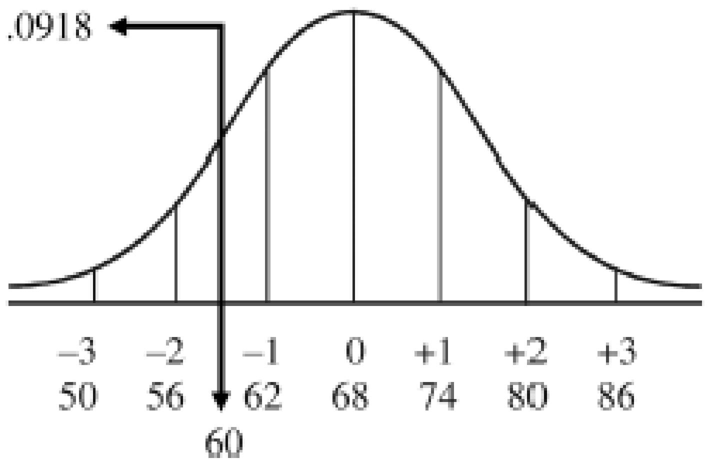
$$p(X > 58) = .9525$$



2. What *proportion* of the population is expected to score less than 60?

$$z = \frac{60 - 68}{6} = -1.33$$

$$p(X < 60) = .0918$$



III. Probabilities/Proportions/Percentages for Specified Raw Scores

A normally distributed set of scores has a $\mu = 80$ and a $\sigma = 12$.

A. What *proportion* of the population can be expected to score above 86?

____

B. What is the *probability* of obtaining a score less than 95? __________

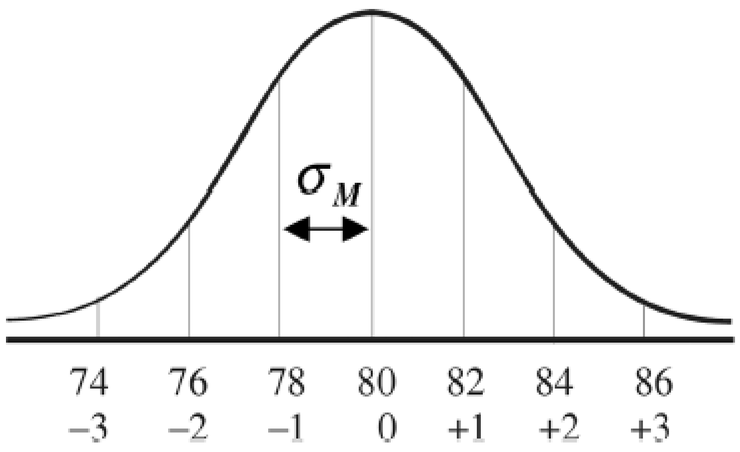
PROBABILITIES, PROPORTIONS, AND PERCENTAGES OF SAMPLE MEANS

The sampling distribution of means can be used to determine the probabilities, proportions, and percentages associated with particular sample means.

$$z = \frac{M - \mu}{\sigma_M} \qquad \text{and} \qquad M = \mu + (z)(\sigma_M)$$

μ = 80, σ = 14, and n = 49. Using the appropriate formula, we find the standard error (σM) to be 2. Thus,

$$\sigma_M = \frac{14}{\sqrt{49}} = 2$$



Given the normally shaped distribution with $\mu = 80$ *and* $\sigma_m = 2$, *A. What is the probability that an obtained sample mean will be below 81?*

What proportion of the sample means can be expected to have a value greater than 83?

- Convert the given sample mean to a z-score:

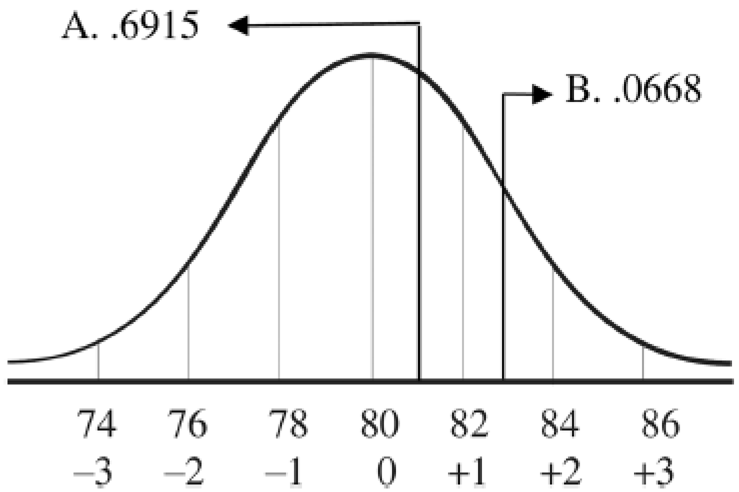
$$z = \frac{M-\mu}{\sigma_M} = \frac{81-80}{2} = +.50$$

- Then, locate in the z-table the probability associated with sample means below a z-score of +.50 (column B):

$$p(M < 81) = .6915$$

$$z = \frac{M - \mu}{\sigma_M} = \frac{83 - 80}{2} = +1.50$$

$$p(M < 83) = .0668$$



Given a normally shaped population distribution with μ =
95 and σ = 5, a sample size of n = 25 is drawn at random.

A. The probability is .05 that the mean of the sample will be above what value?

- First, find the standard error: $\sigma_M = \dfrac{5}{\sqrt{25}} = 1$
- Next, find the z-score associated with the closest proportion above .0500, which is 1.65.
- Finally, use the formula to convert the z-score to a sample mean:

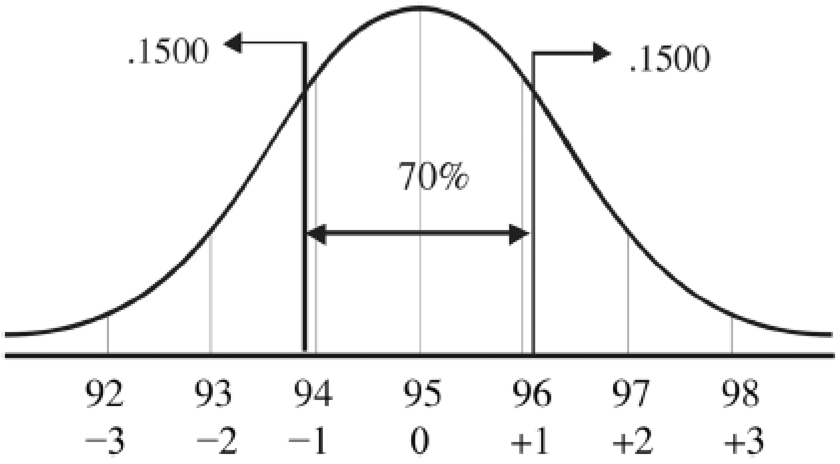
$$\begin{aligned} M &= \mu + (z)(\sigma_M) \\ &= 95 + (+1.65)(1) \\ &= 96.65 \end{aligned}$$

What range of sample means would be expected to occur in the middle of the distribution 70% of the time?

$$M = \mu + (z)(\sigma_M)$$

$$M = 95 + (-1.04)(1) \qquad M = 95 + (+1.04)(1)$$

$$= 93.96 \qquad\qquad = 96.04$$



Column A	Column B	Column C	Column D	Column A	Column B	Column C	Column D
z-score	Proportion in body (larger part)	Proportion in tail (smaller part)	Proportion between mean and *z*	*z*-score	Proportion in body (larger part)	Proportion in tail (smaller part)	Proportion between mean and *z*
0.96	.8315	.1685	.3315	**1.45**	.9265	.0735	.4265
0.97	.8340	.1660	.3340	**1.46**	.9279	.0721	.4279
0.98	.8365	.1635	.3365	**1.47**	.9292	.0708	.4292
0.99	.8389	.1611	.3389	**1.48**	.9306	.0694	.4306
1.00	.8413	.1587	.3413	**1.49**	.9319	.0681	.4319
1.01	.8438	.1562	.3438	**1.50**	.9332	.0668	.4332
1.02	.8461	.1539	.3461	**1.51**	.9345	.0655	.4345
1.03	.8485	.1515	.3485	**1.52**	.9357	.0643	.4357
1.04	.8508	.1492	.3508	**1.53**	.9370	.0630	.4370
1.05	.8531	.1469	.3531	**1.54**	.9382	.0618	.4382
1.06	.8554	.1446	.3554	**1.55**	.9394	.0606	.4394
1.07	.8577	.1423	.3577	**1.56**	.9406	.0594	.4406

I. Finding Probability or Proportion of Given Sample Means

Given a normal distribution with $\mu = 100$ and $\sigma = 12$, a sample of $n = 36$ is drawn at random.

A. What is the probability that the sample mean will fall above 99? ______________

B. What proportion of the sample means will have a value less than 95? ______________

II. Finding Sample Means from Given Probabilities or Percentages

The Wechsler IQ test is normally distributed and has a known $\mu = 100$ and a $\sigma = 15$. A sample of $n = 25$ is drawn at random.

A. The probability is .04 that the mean of the sample will be below what value?

____

B. What middle range of sample IQs will be expected to occur 95% of the time?

____

Answers to "Your Turn!" Problems

I. Finding Probability or Proportion of Given Sample Means

A. $\sigma_M = \dfrac{\sigma}{\sqrt{n}} = \dfrac{12}{\sqrt{36}} = 2 \quad z = \dfrac{M-\mu}{\sigma_M} = \dfrac{99-100}{2} = -.50 \quad p(M > 99) = .6915$

B. $z = \dfrac{M-\mu}{\sigma_M} = \dfrac{95-100}{2} = -2.50 \quad p(M < 95) = .0062$

Answers to "Your Turn!" Problems
(continued)

II. Finding Sample Means from Given Probabilities or Percentages

A. $\sigma_M = \dfrac{\sigma}{\sqrt{n}} = \dfrac{15}{\sqrt{25}} = 3$

$$\begin{aligned} M &= \mu + (z)(\sigma_M) \\ &= 100 + (-1.75)(3) \\ &= 94.75 \end{aligned}$$

B. Need to identify the extreme 5%, 2½% at each end. The *z*-score associated with .0250 is ±1.96. Thus,

$$\begin{aligned} M &= \mu + (z)(\sigma_M) \\ &= 100 + (-1.96)(3) \\ &= 94.12 \end{aligned}$$

$$\begin{aligned} M &= \mu + (z)(\sigma_M) \\ &= 100 + (+1.96)(3) \\ &= 105.88 \end{aligned}$$