Review of Statistics: Central Limit Theorem

Dr. Patrick Toche

Textbook:

James H. Stock and Mark W. Watson, *Introduction to Econometrics*, 4th Edition, Pearson.

Other references:

Joshua D. Angrist and Jörn-Steffen Pischke, *Mostly Harmless Econometrics: An Empiricist's Companion*, 1st Edition, Princeton University Press.
Jeffrey M. Wooldridge, *Introductory Econometrics: A Modern Approach*, 7th Edition, Cengage Learning.

The textbook comes with online resources and study guides. Other references will be given from time to time.

Central Limit Theorem

In a population, $\mu = 100$ and $\sigma^2 = 43$. Use the central limit theorem to answer the question:

a. In a random sample of size $n = 100$, find $\Pr(\overline{Y} < 101)$.

$$
\begin{aligned}
\mathrm{P}\left[\overline{Y} < 101\right] &= \mathrm{P}\left[\frac{\overline{Y} - \mu}{\sigma/\sqrt{n}} < \frac{101 - \mu}{\sigma/\sqrt{n}}\right] \\
&= \mathrm{P}\left[Z < \frac{101 - 100}{\sqrt{43}/\sqrt{100}}\right] \\
&\approx \mathrm{P}[Z < 1.525] \\
&\approx 0.94
\end{aligned}
$$

In R, you would compute the probability with `pnorm(1.525)`.

Central Limit Theorem

In a population, $\mu = 100$ and $\sigma^2 = 43$. Use the central limit theorem to answer the question:

a. In a random sample of size $n = 100$, find $\Pr(\overline{Y} < 101)$.

$$\begin{aligned}
\mathrm{P}\left[\overline{Y} < 101\right] &= \mathrm{P}\left[\frac{\overline{Y} - \mu}{\sigma/\sqrt{n}} < \frac{101 - \mu}{\sigma/\sqrt{n}}\right] \\
&= \mathrm{P}\left[Z < \frac{101 - 100}{\sqrt{43}/\sqrt{100}}\right] \\
&\approx \mathrm{P}[Z < 1.525] \\
&\approx 0.94
\end{aligned}$$

In R, you would compute the probability with `pnorm(1.525)`.

Central Limit Theorem

In a population, $\mu = 100$ and $\sigma^2 = 43$. Use the central limit theorem to answer the question:

b. In a random sample of size $n = 165$, find $\Pr(\overline{Y} > 98)$.

$$
\begin{aligned}
\mathrm{P}\left[\overline{Y} > 98\right] &= 1 - \mathrm{P}\left[\overline{Y} < 98\right] \\
&= 1 - \mathrm{P}\left[\frac{\overline{Y} - \mu}{\sigma/\sqrt{n}} < \frac{98 - \mu}{\sigma/\sqrt{n}}\right] \\
&= 1 - \mathrm{P}\left[Z < \frac{98 - 100}{\sqrt{43}/\sqrt{165}}\right] \\
&\approx 1 - \mathrm{P}[Z < -3.92] \\
&\approx 0.99995 \\
&\approx 1
\end{aligned}
$$

R code: `1-pnorm(-3.9178)`.

Central Limit Theorem

In a population, $\mu = 100$ and $\sigma^2 = 43$. Use the central limit theorem to answer the question:

b. In a random sample of size $n = 165$, find $\Pr(\overline{Y} > 98)$.

$$
\begin{aligned}
\mathrm{P}\left[\overline{Y} > 98\right] &= 1 - \mathrm{P}\left[\overline{Y} < 98\right] \\
&= 1 - \mathrm{P}\left[\frac{\overline{Y} - \mu}{\sigma/\sqrt{n}} < \frac{98 - \mu}{\sigma/\sqrt{n}}\right] \\
&= 1 - \mathrm{P}\left[Z < \frac{98 - 100}{\sqrt{43}/\sqrt{165}}\right] \\
&\approx 1 - \mathrm{P}[Z < -3.92] \\
&\approx 0.99995 \\
&\approx 1
\end{aligned}
$$

R code: `1-pnorm(-3.9178)`.

Central Limit Theorem

In a population, $\mu = 100$ and $\sigma^2 = 43$. Use the central limit theorem to answer the question:

c. In a random sample of size $n = 64$, find $\Pr(101 < \overline{Y} < 103)$.

$$\begin{aligned}
\mathrm{P}\left[101 < \overline{Y} < 103\right] &= \mathrm{P}\left[\frac{101-\mu}{\sigma/\sqrt{n}} < \frac{\overline{Y}-\mu}{\sigma/\sqrt{n}} < \frac{103-\mu}{\sigma/\sqrt{n}}\right] \\
&= \mathrm{P}\left[\frac{101-100}{\sqrt{43}/\sqrt{64}} < Z < \frac{103-100}{\sqrt{43}/\sqrt{64}}\right] \\
&\approx \mathrm{P}[1.220 < Z < 3.660] \\
&\approx 1 - 0.89 \\
&\approx 0.11
\end{aligned}$$

R code: `pnorm(3.660)-pnorm(1.220)`.

Central Limit Theorem

In a population, $\mu = 100$ and $\sigma^2 = 43$. Use the central limit theorem to answer the question:

c. In a random sample of size $n = 64$, find Pr$(101 < \overline{Y} < 103)$.

$$\begin{aligned}
\mathrm{P}\left[101 < \overline{Y} < 103\right] &= \mathrm{P}\left[\frac{101-\mu}{\sigma/\sqrt{n}} < \frac{\overline{Y}-\mu}{\sigma/\sqrt{n}} < \frac{103-\mu}{\sigma/\sqrt{n}}\right] \\
&= \mathrm{P}\left[\frac{101-100}{\sqrt{43}/\sqrt{64}} < Z < \frac{103-100}{\sqrt{43}/\sqrt{64}}\right] \\
&\approx \mathrm{P}[1.220 < Z < 3.660] \\
&\approx 1 - 0.89 \\
&\approx 0.11
\end{aligned}$$

R code: `pnorm(3.660)-pnorm(1.220)`.