STAT 301: Introduction to Sampling Distributions
Material covered: Sections 7.1, 7.2 of text

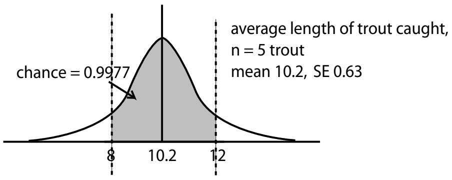
Exercise 7.1 (Sampling distribution $\bar{x}$, normal)

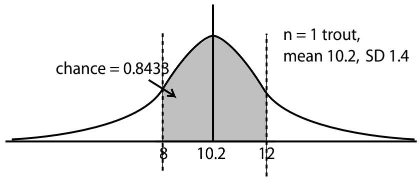
1. *Trout (Example 2, p 300).* Suppose length of trout, x, has a *normal* distribution with mean $\mu = 10.7$ inches and standard deviation of $\sigma = 1.4$ inches.

   (a) *Review: single trout.* Since (nonstandard normal) x can be transformed to *standard* normal,
   $$z = \frac{x - \mu}{\sigma},$$
   chance a *single* ($n = 1$) trout is between 8 and 12 inches long is
   $$\begin{aligned} P(8 < x < 12) &= P\left(\frac{8 - 10.2}{1.4} < z < \frac{12 - 10.2}{1.4}\right) \\ &\approx P(-1.57 < z < 1.29) \approx 0.9015 - 0.0582 = 0.8433. \end{aligned}$$
   (or, using TI–84+: 2nd DISTR normalcdf$(-1.57, 1.29, 0, 1) \approx 0.8433$)

   (b) *New: multiple trout.* Since *average* $\bar{x}$ is *also* (nonstandard) normal and so
   $$\mu_{\bar{x}} = \mu, \quad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \quad \text{(and so)} \quad z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}},$$
   chance *average* of *five* ($n = 5$) trout is between 8 and 12 inches long is
   $$\begin{aligned} P(8 < \bar{x} < 12) &= P\left(\frac{8 - 10.2}{\frac{1.4}{\sqrt{3}}} < z < \frac{12 - 10.2}{\frac{1.4}{\sqrt{3}}}\right) \\ &\approx P\left(\frac{8 - 10.2}{0.63} < z < \frac{12 - 10.2}{0.63}\right) \\ &\approx P(-3.49 < z < 2.86) \approx 0.9979 - 0.0002 = 0.9977. \end{aligned}$$
   (or, using TI–84+: 2nd DISTR normalcdf$(-3.49, 2.86, 0, 1) \approx 0.9976$)

   (c) *Why is $P(8 < \bar{x} < 12) \leq P(8 < x < 12)$?* Because *standard error* $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{1.4}{\sqrt{5}} \approx 0.63$ is less than *standard deviation* $\sigma = 1.4$.





Exercise 7.2 (Sampling distribution $\bar{x}$, Central Limit Theorem)

1. *Milk (Example 3, p 303).* Suppose bacteria count per milliter of milk, x, has an *unknown* distribution with mean $\mu = 2500$ and standard deviation of $\sigma = 300$.

   (a) *Review: one milliter.* What is the chance a *single* ($n = 1$) milliter of milk has bacteria count between 2350 and 2650?
   Answer: Impossible to tell because probability distribution is unknown.

   (b) *New: multiple ($n \geq 30$) milliters.* According to *Central Limit Theorem*, for $n \geq 30$, the sample average, $\bar{x}$ is approximately normal where

   $$\mu_{\bar{x}} = \mu, \quad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \quad \text{(and so)} \quad z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}.$$

   So, chance *average* of $n = 42$ milliters of milk has bacteria count between 2350 and 2650 is

   $$\begin{aligned} P(2350 < \bar{x} < 2650) &\approx P\left(\frac{2350 - 2500}{\frac{300}{\sqrt{42}}} < z < \frac{2650 - 2500}{\frac{300}{\sqrt{42}}}\right) \\ &= P\left(\frac{2350 - 2500}{46.3} < z < \frac{2650 - 2500}{46.3}\right) \\ &\approx P\left(-3.24 < z < 3.24\right) \approx 0.9994 - 0.0006 = 0.9988. \end{aligned}$$

   (or, using TI–84+: 2nd DISTR normalcdf($-3.24, 3.24, 0, 1$) ≈ 0.9988)

   (c) *Another example.* Chance *average* of $n = 42$ milliters of milk has bacteria count between 2450 and 2600 is

   $$\begin{aligned} P(2450 < \bar{x} < 2600) &\approx P\left(\frac{2450 - 2500}{\frac{300}{\sqrt{42}}} < z < \frac{2600 - 2500}{\frac{300}{\sqrt{42}}}\right) \\ &= P\left(\frac{2450 - 2500}{46.3} < z < \frac{2600 - 2500}{46.3}\right) \\ &\approx P\left(-1.08 < z < 2.16\right) \approx 0.9846 - 0.1401 = 0.8445. \end{aligned}$$

   (or, using TI–84+: 2nd DISTR normalcdf($-1.08, 2.16, 0, 1$) ≈ 0.8445)

2. *Tunnel (Guided Exercise 3, p 305).* Suppose time a car travels a tunnel, x, has an *unknown* distribution with mean $\mu = 12.1$ minutes and $\sigma = 3.8$ minutes.

   (a) Chance average of $n = 50$ cars take between 11 and 13 minutes is

   $$\begin{aligned} P(11 < \bar{x} < 13) &\approx P\left(\frac{11 - 12.1}{\frac{3.8}{\sqrt{50}}} < z < \frac{13 - 12.1}{\frac{3.8}{\sqrt{50}}}\right) \\ &= P\left(\frac{11 - 12.1}{0.54} < z < \frac{13 - 12.1}{0.54}\right) \\ &\approx P\left(-2.07 < z < 1.67\right) \approx 0.9525 - 0.0207 = 0.9318. \end{aligned}$$

   (or, using TI–84+: 2nd DISTR normalcdf($-2.07, 1.67, 0, 1$) ≈ 0.9319)

(b) Chance average of $n = 50$ cars take between 11.5 and 13.2 minutes is

$$\begin{aligned} P(11.5 < \bar{x} < 13.2) &\approx P\left(\frac{11.5 - 12.1}{\frac{3.8}{\sqrt{50}}} < z < \frac{13.2 - 12.1}{\frac{3.8}{\sqrt{50}}}\right) \\ &= P\left(\frac{11.5 - 12.1}{0.54} < z < \frac{13.2 - 12.1}{0.54}\right) \\ &\approx P\left(-1.11 < z < 2.04\right) \approx ? \end{aligned}$$

(or, using TI–84+: 2nd DISTR normalcdf$(-1.11, 2.04, 0, 1) \approx 0.8458$)

(c) Chance average of $n = 50$ cars take less than 11.5 minutes is

$$\begin{aligned} P(\bar{x} < 11.5) &\approx P\left(z < \frac{11.5 - 12.1}{\frac{3.8}{\sqrt{50}}}\right) \\ &= P\left(z < \frac{11.5 - 12.1}{0.54}\right) \\ &\approx P\left(z < -1.11\right) \approx ? \end{aligned}$$

(or, using TI–84+: 2nd DISTR normalcdf$(-E99, -1.11, 0, 1) \approx 0.1335$)

(d) What is chance average of $n = 15$ cars take less than 11.5 minutes?

Exercise 7.3 (Definitions)

Sample:	a selected, often random, subset of a population
Statistic:	a numerical quantity, such as an average, $\bar{x}$, standard deviation, s, variance, s^2, or proportion, $\hat{p}$, calculated from a sample
Population:	a set of measurements or observations of any finite or infinite collection of objects
Parameter:	a numerical quantity, such as an average, μ, standard deviation, σ, variance, σ^2, or proportion, p, calculated from a population or of a probability model describing a population
Sampling Distribution:	probability distribution of statistic based on a random sample

1. *Milk (Example 3, p 303).* Suppose bacteria count per milliter of milk, x, has unknown population probability model with mean $\mu = 2500$ and standard deviation of $\sigma = 300$. We are interested in calculating probabilities for sampling distribution of average bacteria count, $\bar{x}$, for $n = 42$ milliters. Match statistical items with appropriate parts of example.

terms	corn example
(a) sample	**(a)** average bacteria count in $n = 42$ milliters, $\bar{x}$
(b) statistic	**(b)** average bacteria count in unknown probability model, μ
(c) population	**(c)** bacteria counts modelled by unknown probability model
(d) parameter	**(d)** bacteria counts for $n = 42$ milliters

terms	(a)	(b)	(c)	(d)
corn example				

2. *Tunnel (Guided Exercise 3, p 305).* Suppose time a car travels a tunnel, X, has unknown population probability model with mean $\mu = 12.1$ minutes and $\sigma = 3.8$ minutes. We are interested in calculating probabilities for sampling distribution of average travel time, $\bar{x}$, for $n = 50$ cars.

terms	corn example
(a) sample	**(a)** travel times modelled by unknown probability model
(b) statistic	**(b)** average travel time in unknown probability model, μ
(c) population	**(c)** average travel time for $n = 50$ cars, $\bar{x}$
(d) parameter	**(d)** travel times for $n = 42$ cars

terms	(a)	(b)	(c)	(d)
corn example				