7.3 The Central Limit Theorem

Let $Y_1, Y_2, \ldots, Y_n$ be independent and identically distributed random variables with $E(Y_i) = \mu$, and $V(Y_i) = \sigma^2 < \infty$. Then for

$$U_n = \frac{\sum_{i=1}^{n} Y_i - n\mu}{\sigma\sqrt{n}} = \frac{\bar{Y} - \mu}{\frac{\sigma}{\sqrt{n}}},$$

$$\lim_{n\to\infty} P\left(U_n \leq u\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-t^2/2}\, dt, \quad \forall u$$

In other words, for large enough n (typically $n > 30$) statistic U_n has approximately a standard normal distribution.

Exercise 7.3 (The Central Limit Theorem)

1. *Burgers.* Suppose number of burgers, Y, made per minute at Best Burger averages $\mu_Y = 2.7$ burgers with standard deviation of $\sigma_Y = 0.64$ of a burger. A number of minutes, in fact $n = 35$ minutes, are chosen at random throughout a typical work day and number of burgers made are counted at these times.

   (a) Expected value of average number of burgers per minute in 35 minutes is $\mu_{\bar{Y}} = \mu_Y =$ (choose one) (i) **2.7** (ii) **2.8** (iii) **2.9** (iv) **3.0**

   (b) Standard deviation of average number of burgers per minute in 35 minutes $\sigma_{\bar{Y}} = \frac{\sigma_Y}{\sqrt{n}} = \frac{0.64}{\sqrt{35}} \approx$ (i) **0.11** (ii) **0.12** (iii) **0.13** (iv) **0.14**

   (c) Chance *average* number of burgers per minute 35 minutes exceeds 2.75

   $$P\left(\bar{Y} > 2.75\right) \approx P\left(Z > \frac{2.75 - \mu}{\frac{\sigma}{\sqrt{n}}}\right) = P\left(Z > \frac{2.75 - 2.7}{\frac{0.64}{\sqrt{35}}}\right) \approx$$

   (choose one) (i) **0.30** (ii) **0.32** (iii) **0.35** (iv) **0.38**
   normalcdf(2.75,E99,2.7,$\frac{0.64}{\sqrt{35}}$) or normalcdf($\frac{2.75-2.7}{\frac{0.64}{\sqrt{35}}}$,E99,0,1)

   (d) Values a and b such that $P\left(a < \bar{Y} < b\right) = 0.95$ given by

   $$P\left(a < \bar{Y} < b\right) \approx P\left(\frac{a-\mu}{\frac{\sigma}{\sqrt{n}}} < Z < \frac{b-\mu}{\frac{\sigma}{\sqrt{n}}}\right) = P\left(\phi_{0.025} < Z < \phi_{0.975}\right) = 0.95$$

   where 97.5th percentile is given by $\frac{b-\mu}{\frac{\sigma}{\sqrt{n}}} = \phi_{0.975} \approx$
   (choose one) (i) **1.28** (ii) **1.645** (iii) **1.96** (iv) **2.58**
   TI-84+: invNorm(0.975)
   and so

   $$b = \phi_{0.975}\left(\frac{\sigma}{\sqrt{n}}\right) + \mu = 1.96\left(\frac{0.64}{\sqrt{35}}\right) + 2.7 \approx$$

(choose one) (i) **2.91** (ii) **2.93** (iii) **2.95** (iv) **2.98**
Also, since $\frac{a-\mu}{\frac{\sigma}{\sqrt{n}}} = \phi_{0.025} \approx -1.96$,

$$a = \phi_{0.025}\left(\frac{\sigma}{\sqrt{n}}\right) + \mu = -1.96\left(\frac{0.64}{\sqrt{35}}\right) + 2.7 \approx$$

(choose one) (i) **2.39** (ii) **2.43** (iii) **2.45** (iv) **2.49**

(e) Chance *total* number of burgers made in 35 minutes exceeds 100

$$\begin{aligned} P\left(\sum Y > 100\right) &= P\left(\bar{Y} > \frac{100}{35}\right) \approx P\left(Z > \frac{\frac{100}{35} - \mu}{\frac{\sigma}{\sqrt{n}}}\right) \\ &= P\left(Z > \frac{\frac{100}{35} - 2.7}{\frac{0.64}{\sqrt{35}}}\right) \approx \end{aligned}$$

(choose one) (i) **0.053** (ii) **0.063** (iii) **0.073** (iv) **0.083**
normalcdf($\frac{\frac{100}{35}-2.7}{\frac{0.64}{\sqrt{35}}}$,E99,0,1)

(f) How many minutes, n, need be sampled so that difference between sample mean, $\bar{Y}$ and population mean, μ, is less than 0.2, with probability 0.95?

$$P\left(\left|\bar{Y} - \mu\right| \leq 0.2\right) = P\left(-\frac{0.2}{\frac{\sigma}{\sqrt{n}}} < Z < \frac{0.2}{\frac{\sigma}{\sqrt{n}}}\right) = P\left(\phi_{0.025} < Z < \phi_{0.975}\right) = 0.95,$$

where 97.5th percentile is given by $\frac{0.2}{\frac{\sigma}{\sqrt{n}}} = \phi_{0.975} \approx$
(choose one) (i) **1.28** (ii) **1.645** (iii) **1.96** (iv) **2.58**
TI-84+: invNorm(0.975)
and so

$$n = \left(\frac{\phi_{0.975}\,\sigma}{0.2}\right)^2 \approx \left(\frac{1.96 \times 0.64}{0.2}\right)^2 \approx$$

(choose one) (i) **37.34** (ii) **38.34** (iii) **39.34** (iv) **40.34**
or 40 minutes.

2. *Temperatures.* Suppose temperature, Y, on any given day during winter in Laporte averages $\mu_Y = 0$ degrees with standard deviation $\sigma_Y = 1$ degree. Temperature is measured on forty ($n = 40$) days; each day is chosen at random.

(a) Expected value of average temperature over 40 days is
$\mu_{\bar{Y}} = \mu_Y =$ (choose one) (i) **−0.5** (ii) **0** (iii) **0.5** (iv) **1**

(b) Standard deviation of average temperature over 40 days
$\sigma_{\bar{Y}} = \frac{\sigma_Y}{\sqrt{n}} = \frac{1}{\sqrt{40}} \approx$ (i) **0.128** (ii) **0.138** (iii) **0.148** (iv) **0.158**

(c) Chance *average* temperature, over 40 days, less than 0.05 is

$$P\left(\bar{Y} < 0.05\right) \approx P\left(Z < \frac{0.05 - \mu}{\frac{\sigma}{\sqrt{n}}}\right) = P\left(Z > \frac{0.05 - 0}{\frac{1}{\sqrt{40}}}\right) \approx$$

(choose one) (i) **0.524** (ii) **0.624** (iii) **0.724** (iv) **0.824**
normalcdf($-E99$, $\frac{0.05-0}{\frac{1}{\sqrt{40}}}$,0,1)

Notice chance temperature less than 0.05
$P(Y < 0.05) \approx$ (i) **0.520** (ii) **0.620** (iii) **0.720** (iv) **0.820**
(2nd DISTR normalcdf($-E99, 0.05, 0, 1$))
(choose one) (i) **does** (ii) **does not** equal
chance *average* temperature is less than 0.05, $P(\bar{Y} < 0.05) \approx 0.624$.

(d) Values a and b such that $P\left(a < \bar{Y} < b\right) = 0.90$ given by

$$P\left(a < \bar{Y} < b\right) \approx P\left(\frac{a - \mu}{\frac{\sigma}{\sqrt{n}}} < Z < \frac{b - \mu}{\frac{\sigma}{\sqrt{n}}}\right) = P\left(\phi_{0.05} < Z < \phi_{0.95}\right) = 0.90$$

where 95th percentile is given by $\frac{b-\mu}{\frac{\sigma}{\sqrt{n}}} = \phi_{0.95} \approx$
(choose one) (i) **1.28** (ii) **1.645** (iii) **1.96** (iv) **2.58**
TI-84+: invNorm(0.95)
and so

$$b = \phi_{0.95}\left(\frac{\sigma}{\sqrt{n}}\right) + \mu = 1.645\left(\frac{1}{\sqrt{40}}\right) + 0 \approx$$

(choose one) (i) **0.16** (ii) **0.26** (iii) **0.36** (iv) **0.46**
Also, since $\frac{a-\mu}{\frac{\sigma}{\sqrt{n}}} = \phi_{0.05} \approx -1.645$,

$$a = \phi_{0.05}\left(\frac{\sigma}{\sqrt{n}}\right) + \mu = -1.645\left(\frac{1}{\sqrt{40}}\right) + 0 \approx$$

(choose one) (i) **−0.26** (ii) **−0.16** (iii) **0.16** (iv) **0.26**

(e) How many days, n, need be sampled so that difference between sample mean, $\bar{Y}$ and population mean, μ, is less than 0.3, with probability 0.90?

$$P\left(\left|\bar{Y} - \mu\right| \leq 0.3\right) = P\left(-\frac{0.3}{\frac{\sigma}{\sqrt{n}}} < Z < \frac{0.3}{\frac{\sigma}{\sqrt{n}}}\right) = P\left(\phi_{0.05} < Z < \phi_{0.95}\right) = 0.90,$$

where 95th percentile is given by $\frac{0.02}{\frac{\sigma}{\sqrt{n}}} = \phi_{0.95} \approx$
(choose one) (i) **1.28** (ii) **1.645** (iii) **1.96** (iv) **2.58**
TI-84+: invNorm(0.95)
and so

$$n = \left(\frac{\phi_{0.95}\,\sigma}{0.3}\right)^2 \approx \left(\frac{1.645 \times 1}{0.3}\right)^2 \approx$$

(choose one) (i) **30.06** (ii) **33.34** (iii) **35.34** (iv) **36.34**
or 31 days.

3. *Temperatures in Summer and Winter.* Suppose temperature, X, on any given day during *summer* averages $\mu_1 = 65$ degrees with *variance* $\sigma_1^2 = 5$ degrees and is measured on thirty-five ($n_1 = 35$) randomly chosen days. Further suppose temperature, Y, on any given day during *winter* averages $\mu_2 = 30$ degrees with variance $\sigma_2^2 = 3$ degrees and is measured on forty ($n_2 = 40$) randomly chosen days.

   (a) Expected value of difference in average temperatures summer and winter
   $\mu_{\bar{X}-\bar{Y}} = \mu_1 - \mu_2 =$ (choose one) (i) **25** (ii) **30** (iii) **35** (iv) **40**

   (b) SD of difference in average temperatures summer and winter
   $\sigma_{\bar{X}-\bar{Y}} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{\frac{5}{35} + \frac{3}{40}} \approx$
   (choose one) (i) **0.367** (ii) **0.467** (iii) **0.567** (iv) **0.667**

   (c) Chance difference in average temperatures less than 34^o

$$P\left(\bar{X} - \bar{Y} \leq 34\right) = P\left(Z < \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}\right)$$
$$= P\left(Z < \frac{34 - 35}{\sqrt{\frac{5}{35} + \frac{3}{40}}}\right) \approx$$

   (choose one) (i) **0.016** (ii) **0.026** (iii) **0.036** (iv) **0.046**
   $\text{normalcdf}\left(-E99, \frac{34-35}{\sqrt{\frac{5}{35}+\frac{3}{40}}}, 0, 1\right)$

4. *Importance of Central Limit Theorem.* The CLT useful because (circle none, one or more):

   (a) No matter what original parent distribution is, as long as a large enough random sample is taken, average of this sample follows a normal (not a binomial or any other distribution) distribution.

   (b) In practical situations where it is not known what parent probability distribution to use, as long as a large enough random sample is taken, average of this sample follows a normal distribution.

   (c) Rather than having to deal with many different probability distributions, as long as a large enough random sample is taken, average of this sample follows *one* distribution, normal distribution.

   (d) Many distributions in statistics rely in one way or another on normal distribution because of CLT.

7.4 A Proof of the Central Limit Theorem (Optional)

Not covered.

7.5 The Normal Approximation to the Binomial Distribution

Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed random variables, where, in particular,

$$X_i = \begin{cases} 1, & i\text{th trial is a success} \\ 0 & \text{otherwise} \end{cases}$$

and let $Y = \sum_{i=1}^{n} X_i$, and so $E(Y) = p$, $V(Y) = \sigma^2 = p(1-p) = pq$. Then for

$$\begin{aligned} U_n &= \frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}} = \frac{Y - np}{\sqrt{npq}} \\ &= \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{\frac{Y}{n} - p}{\sqrt{\frac{pq}{n}}} \end{aligned}$$

$$\lim_{n \to \infty} P(U_n \leq u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-t^2/2}\, dt, \quad \forall u$$

In this version of the Central Limit Theorem, X_i are bernoulli and Y is binomial.

Exercise 7.5 (The Normal Approximation to the Binomial Distribution)

1. *Defective Widgets.* Each of $n = 75$ widgets are defective 55% ($p = 0.55$) of the time. Assume this problem obeys conditions of a binomial experiment.

   (a) *(exact) binomial:* Chance at most 38 widgets defective
   $P(Y \leq 38) \approx$ (choose one) (i) **0.261** (ii) **0.361** (iii) **0.461** (iv) **0.561**
   DISTR A:binomcdf(75, 0.55, 38).

   (approximate) normal with continuity correction:
   Chance at most 38 widgets defective

   $$P(Y \leq 38) \approx P\left(Z \leq \frac{38.5 - np}{\sqrt{npq}}\right) = P\left(Z \leq \frac{38.5 - 75(0.55)}{\sqrt{75(0.55)(0.45)}}\right) \approx$$

   (choose one) (i) **0.161** (ii) **0.202** (iii) **0.232** (iv) **0.262**
   normalcdf($-E99$, $\frac{38.5-75(0.55)}{\sqrt{75(0.55)(0.45)}}$,0,1)

(b) *(exact) binomial:* Chance between 38 and 43 widgets defective
$P(38 \leq Y \leq 43) \approx$ (i) **0.395** (ii) **0.461** (iii) **0.493** (iv) **0.506**
DISTR binomcdf(75, 0.55, 43) − DISTR binomcdf(75, 0.55, 37).

(approximate) normal with continuity correction:
Chance between 38 and 43 widgets defective

$$P(38 \leq Y \leq 43) \approx P\left(\frac{37.5 - np}{\sqrt{npq}} \leq Z \leq \frac{43.5 - np}{\sqrt{npq}}\right)$$
$$= P\left(\frac{37.5 - 75(0.55)}{\sqrt{75(0.55)(0.45)}} \leq Z \leq \frac{43.5 - 75(0.55)}{\sqrt{75(0.55)(0.45)}}\right) \approx$$

(choose one) (i) **0.423** (ii) **0.459** (iii) **0.507** (iv) **0.534**
normalcdf($\frac{37.5-75(0.55)}{\sqrt{75(0.55)(0.45)}}$, $\frac{43.5-75(0.55)}{\sqrt{75(0.55)(0.45)}}$,0,1)

(c) *(approximate) normal without continuity correction:*
Chance between 0.5 and 0.6 *proportion* defective widgets

$$P\left(0.5 \leq \frac{Y}{n} \leq 0.6\right) \approx P\left(\frac{0.5 - p}{\sqrt{\frac{pq}{n}}} \leq Z \leq \frac{0.6 - p}{\sqrt{\frac{pq}{n}}}\right)$$
$$= P\left(\frac{0.5 - 0.55}{\sqrt{\frac{0.55(0.45)}{75}}} \leq Z \leq \frac{0.6 - 0.55}{\sqrt{\frac{0.55(0.45)}{75}}}\right) \approx$$

(choose one) (i) **0.529** (ii) **0.616** (iii) **0.676** (iv) **0.734**
normalcdf($\frac{0.5-0.55}{\sqrt{\frac{0.55(0.45)}{75}}}$, $\leq$ $\frac{0.6-0.55}{\sqrt{\frac{0.55(0.45)}{75}}}$,0,1)

(approximate) normal with continuity correction:
Chance between 0.5 and 0.6 *proportion* defective widgets

$$P\left(0.5 \leq \frac{Y}{n} \leq 0.6\right) \approx P\left(\frac{0.5 - \frac{0.5}{75} - p}{\sqrt{\frac{pq}{n}}} \leq Z \leq \frac{0.6 + \frac{0.5}{75} - p}{\sqrt{\frac{pq}{n}}}\right)$$
$$= P\left(\frac{0.5 - \frac{0.5}{75} - 0.55}{\sqrt{\frac{0.55(0.45)}{75}}} \leq Z \leq \frac{0.6 + \frac{0.5}{75} - 0.55}{\sqrt{\frac{0.55(0.45)}{75}}}\right) \approx$$

(choose one) (i) **0.529** (ii) **0.616** (iii) **0.676** (iv) **0.734**
normalcdf($\frac{0.5-\frac{0.5}{75}-0.55}{\sqrt{\frac{0.55(0.45)}{75}}}$, $\leq$ $\frac{0.6+\frac{0.5}{75}-0.55}{\sqrt{\frac{0.55(0.45)}{75}}}$,0,1)

(d) How many widgets, n, need be sampled so sample proportion, $\frac{Y}{n}$, is within 0.1 of population proportion, p, with probability 0.99? Do not use continuity correction.

$$P\left(\left|\frac{Y}{n} - p\right| \leq 0.1\right) = P\left(-\frac{0.1}{\sqrt{\frac{pq}{n}}} < Z < \frac{0.1}{\sqrt{\frac{pq}{n}}}\right) = P(\phi_{0.005} < Z < \phi_{0.995}) = 0.99,$$

where 99.5th percentile is given by $\frac{0.1}{\sqrt{\frac{pq}{n}}} = \phi_{0.995} \approx$
(choose one) (i) **1.28** (ii) **1.645** (iii) **1.96** (iv) **2.58**
TI-84+: invNorm(0.995)
and so

$$n = pq\left(\frac{\phi_{0.995}}{0.1}\right)^2 \approx 0.55 \times 0.45 \times \left(\frac{2.58}{0.1}\right)^2 \approx$$

(choose one) (i) **159.7** (ii) **161.7** (iii) **164.7** (iv) **169.7**
or 170 widgets.

(e) Normal approximation to binomial adequate when proportion very much in middle of (0,1); more specifically, $p \pm 3\sqrt{\frac{pq}{n}}$ is in interval $(0, 1)$ or, after some manipulation,

$$n > 9\left(\frac{\max(p,q)}{\min(p,q)}\right).$$

In this case, $n = 75$ and

$$9\left(\frac{\max(p,q)}{\min(p,q)}\right) = 9\left(\frac{\max(0.55, 0.45)}{\min(0.55, 0.45)}\right) =$$

(choose one) (i) **9** (ii) **10** (iii) **11** (iv) **12**
so, since $n = 75 > 11$, normal approximation to binomial in this case is very adequate.

2. *Winning Cases.* A lawyer estimates she wins 40% of her cases ($p = 0.4$), and currently represents $n = 30$ clients. Let Y represent number of wins (of 30 cases) and so $\frac{Y}{n}$ is proportion of wins (of 30 cases). Assume this problem obeys conditions of a binomial experiment.

   (a) *(exact) binomial:* Chance lawyer wins at least 10 cases
   $P(Y \geq 10) \approx$ (choose one) (i) **0.801** (ii) **0.824** (iii) **0.857** (iv) **0.899**
   DISTR 1 - binomcdf(30, 0.4, 9); notice it is "1 - ..." and the "9", not "10"!

   (approximate) normal with continuity correction:
   Chance lawyer wins at least 10 cases

   $$P(Y \geq 10) \approx P\left(Z \geq \frac{9.5 - np}{\sqrt{npq}}\right) = P\left(Z \geq \frac{9.5 - 30(0.4)}{\sqrt{30(0.4)(0.6)}}\right) \approx$$

   (choose one) (i) **0.824** (ii) **0.883** (iii) **0.892** (iv) **0.903**
   normalcdf($\frac{9.5-30(0.4)}{\sqrt{30(0.4)(0.6)}}$, $E99$,0,1)

   (b) *(exact) binomial:* Chance lawyer wins more than 10 cases
   $P(Y > 10) \approx$ (choose one) (i) **0.600** (ii) **0.633** (iii) **0.658** (iv) **0.709**

DISTR 1 - binomcdf(30, 0.4, 10)

(approximate) normal with continuity correction:
Chance lawyer wins more than 10 cases

$$P(Y > 10) \approx P\left(Z > \frac{10.5 - np}{\sqrt{npq}}\right) = P\left(Z > \frac{10.5 - 30(0.4)}{\sqrt{30(0.4)(0.6)}}\right) \approx$$

(choose one) (i) **0.624** (ii) **0.693** (iii) **0.703** (iv) **0.712**
normalcdf($\frac{10.5-30(0.4)}{\sqrt{30(0.4)(0.6)}}$, E99,0,1)

(c) *(approximate) normal without continuity correction:*
Chance between 0.5 and 0.6 *proportion* wins (out of 30)

$$P\left(0.5 \leq \frac{Y}{n} \leq 0.6\right) \approx P\left(\frac{0.5 - p}{\sqrt{\frac{pq}{n}}} \leq Z \leq \frac{0.6 - p}{\sqrt{\frac{pq}{n}}}\right) = P\left(\frac{0.5 - 0.4}{\sqrt{\frac{0.4(0.6)}{30}}} \leq Z \leq \frac{0.6 - 0.4}{\sqrt{\frac{0.4(0.6)}{30}}}\right) \approx$$

(choose one) (i) **0.089** (ii) **0.099** (iii) **0.109** (iv) **0.119**
normalcdf($\frac{0.5-0.4}{\sqrt{\frac{0.4(0.6)}{30}}}$, $\leq\leq$ $\frac{0.6-0.4}{\sqrt{\frac{0.4(0.6)}{30}}}$,0,1)

(approximate) normal with continuity correction:
Chance between 0.5 and 0.6 *proportion* defective widgets

$$P\left(0.5 \leq \frac{Y}{n} \leq 0.6\right) \approx P\left(\frac{0.5 - \frac{0.5}{30} - p}{\sqrt{\frac{pq}{n}}} \leq Z \leq \frac{0.6 + \frac{0.5}{30} - p}{\sqrt{\frac{pq}{n}}}\right) = P\left(\frac{0.5 - \frac{0.5}{30} - 0.4}{\sqrt{\frac{0.4(0.6)}{30}}} \leq Z \leq \frac{0.6 + \frac{0.5}{30} - 0.4}{\sqrt{\frac{0.4(0.6)}{30}}}\right) \approx$$

(choose one) (i) **0.089** (ii) **0.123** (iii) **0.156** (iv) **0.234**
normalcdf($\frac{0.5-\frac{0.5}{30}-0.4}{\sqrt{\frac{0.4(0.6)}{30}}}$, $\leq\leq$ $\frac{0.6+\frac{0.5}{30}-0.4}{\sqrt{\frac{0.4(0.6)}{30}}}$,0,1)

(d) How many clients, n, need be handled by lawyer so sample proportion, $\frac{Y}{n}$, is within 0.1 of population proportion, p, with probability 0.99? Do not use continuity correction.

$$P\left(\left|\frac{Y}{n} - p\right| \leq 0.1\right) = P\left(-\frac{0.1}{\sqrt{\frac{pq}{n}}} < Z < \frac{0.1}{\sqrt{\frac{pq}{n}}}\right) = P(\phi_{0.005} < Z < \phi_{0.995}) = 0.99,$$

where 99.5th percentile is given by $\frac{0.1}{\sqrt{\frac{pq}{n}}} = \phi_{0.995} \approx$
(choose one) (i) **1.28** (ii) **1.645** (iii) **1.96** (iv) **2.58**

TI-84+: invNorm(0.995)
and so

$$n = pq\left(\frac{\phi_{0.995}}{0.1}\right)^2 \approx 0.4 \times 0.6 \times \left(\frac{2.58}{0.1}\right)^2 \approx$$

(choose one) (i) **149.7** (ii) **151.7** (iii) **154.7** (iv) **159.8**
or 160 defendants.

(e) Since $n = 30$ and

$$9\left(\frac{\max(p,q)}{\min(p,q)}\right) = 9\left(\frac{\max(0.4,0.6)}{\min(0.4,0.6)}\right) =$$

(choose one) (i) **11.5** (ii) **13.5** (iii) **15.5** (iv) **16.5**
and so $n = 30 > 13.5$, normal approximation to binomial in this case is very adequate.

3. *Winning Cases, Two Lawyers.* One lawyer estimates she wins 40% of her cases ($p_1 = 0.4$), and currently represents $n_1 = 30$ clients. Another lawyer estimates she wins 55% of her cases ($p_2 = 0.55$) and currently has $n_2 = 50$ clients.

   (a) Expected value of difference in proportion wins two lawyers
   $p_1 - p_2 =$ (choose one) (i) $\mathbf{-0.20}$ (ii) $\mathbf{-0.15}$ (iii) $\mathbf{-0.10}$ (iv) $\mathbf{-0.05}$

   (b) SD of difference in proportion wins two lawyers
   $\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}} = \sqrt{\frac{0.4\times0.6}{30} + \frac{0.55\times0.45}{50}} \approx$
   (choose one) (i) **0.067** (ii) **0.102** (iii) **0.114** (iv) **0.145**

   (c) Probability difference in sample proportion wins is within 0.1 of difference in population proportion wins

$$\begin{aligned} P\left(\left|\frac{Y_1}{n_1} - \frac{Y_2}{n_2} - (p_1 - p_2)\right| < 0.1\right) &= P\left(\frac{-0.1}{\sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}}} < Z < \frac{0.1}{\sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}}}\right) \\ &= P\left(\frac{-0.1}{\sqrt{\frac{0.4\times0.6}{30} + \frac{0.55\times0.45}{50}}} < Z < \frac{0.1}{\sqrt{\frac{0.4\times0.6}{30} + \frac{0.55\times0.45}{50}}}\right) \end{aligned}$$

   (choose one) (i) **0.620** (ii) **0.657** (iii) **0.712** (iv) **0.735**
   normalcdf$\left(\frac{-0.1}{\sqrt{\frac{0.4\times0.6}{30} + \frac{0.55\times0.45}{50}}}, \frac{0.1}{\sqrt{\frac{0.4\times0.6}{30} + \frac{0.55\times0.45}{50}}}, 0, 1\right)$