Practice Quiz Questions 7 (Attendance 14) for Statistics 225
Introduction to Probability Models
Material Covered: Sections 7.3–7.5 of workbook and text

section, pages, questions	
7.3, pp 373–377	(40)
	(42)
	(44)
	(52)
	(56)
	(60)
7.5, pp 382–385	(68)
	(70)
	(74)
	(76)
	(78)
	(84)

(40, pp 373–377)

(a) Hard to tell any differences between two histograms because there are too few simulations.

(b) The 1-sample simulations give a skewed histogram; average-of-3-sample simulations give a mound-shaped histogram.

(c) Average of 1-sample simulations approximate actual population mean, 16.50; standard deviation (SD) of 1-sample simulations approximate actual population SD, 6.03.

Mean of average-of-3-sample simulations approximate actual population mean, 16.50; SD of average-of-3-sample simulations approximate $\frac{6.03}{\sqrt{3}} \approx$ 3.48.

(d) Histogram more normal (mound) shaped for average-of-3-sample simulations histogram than for 1-sample simulations histogram.

(e) Simulated histogram of average-of-n-sample simulations becomes increasingly normal (mound shaped) as n increases and is most normal shaped at $n = 25$.

(42, pp 373–377)

$\mu = 14, \sigma = 2, n = 100$

(a) Probability for $\bar{Y}$

$$P\left(\bar{Y} > 14.5\right) = P\left(Z > \frac{14.5 - \mu}{\frac{\sigma}{\sqrt{n}}}\right) = P\left(Z > \frac{14.5 - 14}{\frac{2}{\sqrt{100}}}\right) \approx 0.0062$$

normalcdf(14.5,E99,14,$\frac{2}{\sqrt{100}}$) or normalcdf($\frac{14.5-14}{\frac{2}{\sqrt{100}}}$,E99,0,1)

(b) Since

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

where 97.5th percentile[1] is given by $\frac{b-\mu}{\frac{\sigma}{\sqrt{n}}} = \phi_{0.975} \approx 1.96$, so

$$b = 1.96\left(\frac{\sigma}{\sqrt{n}}\right) + \mu = 1.96\left(\frac{2}{\sqrt{100}}\right) + 14 \approx 14.392.$$

Also, since $\frac{a-\mu}{\frac{\sigma}{\sqrt{n}}} = \phi_{0.025} \approx -1.96$,

$$a = -1.96\left(\frac{\sigma}{\sqrt{n}}\right) + \mu = -1.96\left(\frac{2}{\sqrt{100}}\right) + 14 \approx 13.608.$$

[1]TI-84+: invNorm(0.975) $\approx$ 1.96

(44, pp 373–377) Since $\sigma = 2.5$, and

$$P\left(\left|\bar{Y} - \mu\right| \leq 0.4\right) = P\left(-\frac{0.4}{\frac{\sigma}{\sqrt{n}}} < Z < \frac{0.4}{\frac{\sigma}{\sqrt{n}}}\right) = P\left(\phi_{0.025} < Z < \phi_{0.975}\right) = 0.95,$$

where 97.5th percentile[2] is given by $\frac{0.4}{\frac{\sigma}{\sqrt{n}}} = \phi_{0.975} \approx 1.96$, so

$$n = \left(\frac{\phi_{0.975}\sigma}{0.4}\right)^2 \approx \left(\frac{1.96 \times 2.5}{0.4}\right)^2 \approx 150.06$$

or 151 men.

(52, pp 373–377)
$\mu = 200, \sigma = 10, n = 25$

(a) Example of $\bar{Y}$:

$$P\left(199 \leq \bar{Y} \leq 202\right) = P\left(\frac{199 - \mu}{\frac{\sigma}{\sqrt{n}}} \leq Z \leq \frac{202 - \mu}{\frac{\sigma}{\sqrt{n}}}\right) = P\left(\frac{199 - 200}{\frac{10}{\sqrt{25}}} \leq Z \leq \frac{202 - 200}{\frac{10}{\sqrt{25}}}\right) \approx 0.532$$

$$\text{normalcdf}\left(\frac{199-200}{\frac{10}{\sqrt{25}}}, \frac{202-200}{\frac{10}{\sqrt{25}}}, 0, 1\right)$$

(b) Example of $\sum Y$:

$$P\left(\sum Y \leq 5100\right) = P\left(\bar{Y} \leq \frac{5100}{25}\right) = P\left(Z \leq \frac{\frac{5100}{25} - \mu}{\frac{\sigma}{\sqrt{n}}}\right) = P\left(Z \leq \frac{\frac{5100}{25} - 200}{\frac{10}{\sqrt{25}}}\right) \approx 0.9772$$

$$\text{normalcdf}\left(-E99, \frac{\frac{5100}{25} - 200}{\frac{10}{\sqrt{25}}}, 0, 1\right)$$

[2]TI-84+: invNorm(0.975) ≈ 1.96

(56, pp 373–377)
Since $\sigma^2 = 4$ or $\sigma = 2$, $n = 50$, and

$$P\left(\sum Y > 200\right) = P\left(\bar{Y} > \frac{200}{50}\right) = P\left(Z > \frac{\frac{200}{50} - \mu}{\frac{\sigma}{\sqrt{n}}}\right) = P\left(Z > \phi_{0.05}\right) = 0.95,$$

where 5th percentile[3] is given by $\frac{\frac{200}{50}-\mu}{\frac{\sigma}{\sqrt{n}}} = \phi_{0.05} \approx -1.645$, so

$$\mu = \frac{200}{50} + \frac{\phi_{0.05}\sigma}{\sqrt{n}} \approx 4 + \frac{-1.645 \times 2}{\sqrt{50}} \approx 4.47.$$

(60, pp 373–377) Since $\sigma_1^2 = 0.01$, $\sigma_2^2 = 0.02$, $n_1 = 50$, and $n_2 = 100$,

$$\begin{aligned} P\left(\left|\bar{X} - \bar{Y} - (\mu_1 - \mu_2)\right| \leq 0.05\right) &= P\left(-\frac{0.05}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} < Z < \frac{0.05}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}\right) \\ &= P\left(-\frac{0.05}{\sqrt{\frac{0.01}{50} + \frac{0.02}{100}}} < Z < \frac{0.05}{\sqrt{\frac{0.01}{50} + \frac{0.02}{100}}}\right) \approx 0.9876. \end{aligned}$$

$$\text{normalcdf}\left(-\frac{0.05}{\sqrt{\frac{0.01}{50}+\frac{0.02}{100}}}, \frac{0.05}{\sqrt{\frac{0.01}{50}+\frac{0.02}{100}}}, 0, 1\right)$$

[3] TI-84+: invNorm(0.05) ≈ -1.645

(68, pp 382–385)
Since $p = 0.48$, $n = 50$,

(a) exact binomial: $P(Y \geq 29) = 1 - P(Y \leq 28) \approx 0.10135$
1 - binomcdf(50,0.48,28)

approximate normal with continuity correction:

$$P(Y \geq 28.5) = P\left(Z \geq \frac{28.5 - np}{\sqrt{np(1-p)}}\right) = P\left(Z \geq \frac{28.5 - 50(0.48)}{\sqrt{50(0.48)(1-0.48)}}\right) \approx 0.10137$$

normalcdf$\left(\frac{28.5-50(0.48)}{\sqrt{50(0.48)(1-0.48)}}, E99, 0, 1\right)$

(b) normal approximation is close to exact binomial because binomial histogram is normal (mound) shaped and because

$$n = 50 > 9 \times \frac{\max(p,q)}{\min(p,q)} = 9 \times \frac{\max(0.48, 0.52)}{\min(0.48, 0.52)} = 9 \times \frac{0.52}{0.42} \approx 9.75$$

(70, pp 382–385)

(a) Since $q = 1 - p$,

$$p + 3\sqrt{\frac{pq}{n}} < 1 \Leftrightarrow \sqrt{\frac{pq}{n}} < \frac{1-p}{3} \Leftrightarrow \frac{pq}{n} < \frac{1}{9}q^2 \Leftrightarrow 9\frac{p}{q^2} < n \Leftrightarrow n > 9\left(\frac{p}{q}\right)$$

(b) similar argument leads to $n > 9\left(\frac{q}{p}\right)$

(c) combining (a) and (b)

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

(74, pp 382–385)

(a) Normal approximation with continuity correction, $p = \frac{1}{410}$, all Americans,

$$P(Y \geq 0.5) = P\left(Z \geq \frac{0.5 - np}{\sqrt{np(1-p)}}\right) = P\left(Z \geq \frac{0.5 - 1500\left(\frac{1}{410}\right)}{\sqrt{1500\left(\frac{1}{410}\right)\left(\frac{409}{410}\right)}}\right) \approx 0.9504$$

$$\text{normalcdf}\left(\frac{0.5-1500(\frac{1}{410})}{\sqrt{1500(\frac{1}{410})(\frac{409}{410})}}, E99, 0, 1\right)$$

(b) Normal approximation continuity correction, $p = \frac{1}{64}$, $n = 1500$, lawyers,

$$P(Y \geq 30.5) = P\left(Z \geq \frac{30.5 - np}{\sqrt{np(1-p)}}\right) = P\left(Z \geq \frac{30.5 - 1500\left(\frac{1}{64}\right)}{\sqrt{1500\left(\frac{1}{64}\right)\left(\frac{63}{64}\right)}}\right) \approx 0.0708$$

$$\text{normalcdf}\left(\frac{30.5-1500(\frac{1}{64})}{\sqrt{1500(\frac{1}{64})(\frac{63}{64})}}, E99, 0, 1\right)$$

(c) Normal approximation continuity correction, $p = \frac{1}{64}$, $n = 1000$, lawyers,

$$P(Y \geq 30.5) = P\left(Z \geq \frac{30.5 - np}{\sqrt{np(1-p)}}\right) = P\left(Z \geq \frac{30.5 - 1000\left(\frac{1}{64}\right)}{\sqrt{1000\left(\frac{1}{64}\right)\left(\frac{63}{64}\right)}}\right) \approx 0$$

$$\text{normalcdf}\left(\frac{30.5-1000(\frac{1}{64})}{\sqrt{1000(\frac{1}{64})(\frac{63}{64})}}, E99, 0, 1\right)$$

Since chance of getting 30 out of 1000 lawyers is so small (unusual), $P(Y \geq 30.5) \approx 0$, this indicates density of lawyers at this corner exceeds density of lawyers in entire city.

(76, pp 382–385)

(a) Variance is

$$V\left(\frac{Y}{n}\right) = \frac{1}{n^2}V(Y) = \frac{np(1-p)}{n^2} = \frac{p-p^2}{n}$$

so

$$\frac{\partial}{\partial p}\left(\frac{p-p^2}{n}\right) = \frac{1-2p}{n} = 1-2p = 0$$

and so maximum occurs at $p = 0.5$.

(b) Since maximum variance occurs at $p = 0.5$,

$$P\left(\left|\frac{Y}{n} - p\right| \leq 0.1\right) = P\left(-\frac{0.1}{\sqrt{\frac{pq}{n}}} < Z < \frac{0.1}{\sqrt{\frac{pq}{n}}}\right) = P\left(\phi_{0.025} < Z < \phi_{0.975}\right) = 0.95,$$

where 97.5th percentile[4] is given by $\frac{0.1}{\sqrt{\frac{pq}{n}}} = \phi_{0.975} \approx 1.96$, so

$$n = pq\left(\frac{\phi_{0.975}}{0.1}\right)^2 \approx 0.5 \times 0.5 \times \left(\frac{1.96}{0.1}\right)^2 \approx 96.04.$$

or 97 items.

(78, pp 382–385)
Since $p = 0.9$, $n = 50$,

$$P\left(\left|\frac{Y}{n} - p\right| \leq 0.15\right) = P\left(-\frac{0.15}{\sqrt{\frac{pq}{n}}} < Z < \frac{0.15}{\sqrt{\frac{pq}{n}}}\right) = P\left(-\frac{0.15}{\sqrt{\frac{0.9(0.1)}{50}}} < Z < \frac{0.15}{\sqrt{\frac{0.9(0.1)}{50}}}\right) \approx 1$$

(84, pp 382–385)

(a) Expected value is

$$E\left(\frac{Y_1}{n_1} - \frac{Y_2}{n_2}\right) = E\left(\frac{Y_1}{n_1}\right) - E\left(\frac{Y_2}{n_2}\right) = \frac{1}{n_1}E(Y_1) - \frac{1}{n_2}E(Y_2) = \frac{n_1p_1}{n_1} - \frac{n_2p_2}{n_2} = p_1 - p_2$$

(b) Variance is

$$\begin{aligned} V\left(\frac{Y_1}{n_1} - \frac{Y_2}{n_2}\right) &= V\left(\frac{Y_1}{n_1}\right) + V\left(\frac{Y_2}{n_2}\right) = \frac{1}{n_1^2}V(Y_1) + \frac{1}{n_2^2}V(Y_2) \\ &= \frac{n_1p_1q_1}{n_1^2} + \frac{n_2p_2q_2}{n_2^2} = \frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2} \end{aligned}$$

[4] TI-84+: invNorm(0.975) $\approx$ 1.96