Chapter 4: Probability and Probability Distributions

4.1
- a. Subjective probability
- b. Relative frequency
- c. Classical
- d. Relative frequency
- e. Subjective probability
- f. Subjective probability
- g. Classical

4.2 Answers will vary depending on person's experience with each situation.

- c. This is know as the "birthday problem." Students will be surprised to learn that in a class of only 23 students the probability of 2 or more students having the same birthday is greater than 50%.

4.3 Using the binomial formula, the probability of guessing correctly 15 or more of the 20 questions is 0.021.

4.4
- a. Positive Outcomes are 00 to 74; Negative Outcomes are 75 to 99.
- b. Let M be the number of sets of 20 2-digit numbers out of the 2000 sets in which we have 15 or more "positive outcomes" (i.e., 15 or more two digit numbers in the range 00 to 74). Then the probability would be approximated by M/2000.

4.5 HHH, HHT, HTH, THH, TTH, THT, HTT, TTT

4.6
- a. A = {HTT, THT, TTH}; Thus, P(A) = $\frac{3}{8}$
- b. B = {HHH, HHT, HTH, THH, THT, HTT, TTH}; Thus, P(B) = $\frac{7}{8}$
- c. C = {TTT}; Thus, P(C) = $\frac{1}{8}$

4.7
- a. $P\left(\overline{A}\right) = 1 - \frac{3}{8} = \frac{5}{8}$
  $P\left(\overline{B}\right) = 1 - \frac{7}{8} = \frac{1}{8}$
  $P\left(\overline{C}\right) = 1 - \frac{1}{8} = \frac{7}{8}$
- b. Events A and B are not mutually exclusive because B contains A $\Rightarrow$ $A \cap B = A$, which is not the empty set.

4.8
- a. $P(A|B) = P(A \cap B)/P(B) = \frac{3}{8}/\frac{7}{8} = \frac{3}{7}$
- b. $P(A|C) = P(A \cap C)/P(C) = 0$, since $A \cap C$ is empty.
- c. $P(B|C) = P(B \cap C)/P(C) = 0$, since $B \cap C$ is empty.

4.9 Because $P(A|B) = \frac{3}{7} \neq \frac{3}{8} = P(A) \Rightarrow$ A and B are not independent.
Because $P(A|C) = 0 \neq \frac{3}{8} = P(A) \Rightarrow$ A and C are not independent.
Because $P(B|C) = 0 \neq \frac{7}{8} = P(B) \Rightarrow$ B and C are not independent.

4.10 A = $\{1, 3, 5\}$ and B = $\{4, 5\}$

a. $P(A) = P(1) + P(3) + P(5) = 0.20 + 0.15 + 0.30 = 0.65$
$P(B) = P(4) + P(5) = 0.10 + 0.30 = 0.40$

b. $P(A \cap B) = P(5) = 0.30$

c. $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.65 + 0.40 - 0.30 = 0.75$
Alternatively, $P(A \cup B) = P(1) + P(3) + P(4) + P(5) = 0.20 + 0.15 + 0.10 + 0.30 = 0.75$

4.11 No, since A and B are not mutually exclusive. Also, $P(A \cap B) = 0.30 \neq 0$

4.12 a. $\overline{A}$: Generator 1 does not work

b. $B|A$: Generator 2 does not work given that Generator 1 does not work

c. $A \cup B$: Generator 1 works or Generator 2 works or Both Generators work

4.13 a. $S = \{F_1F_2, F_1F_3, F_1F_4, F_1F_5, F_2F_1, F_2F_3, F_2F_4, F_2F_5, F_3F_1, F_3F_2,$
$F_3F_4, F_3F_5, F_4F_1, F_4F_2, F_4F_3, F_4F_5, F_5F_1, F_5F_2, F_5F_3, F_5F_4\}$

b. Let T_1 be the event that the 1st firm chosen is stable and T_2 be the event that the 2nd firm chosen is stable.
$P(T_1) = \frac{3}{5}$ and $P(\overline{T_1}) = \frac{2}{5}$
$P(\text{Both Stable}) = P(T_1 \cap T_2) = P(T_2|T_1)P(T_1) = (\frac{2}{4})(\frac{3}{5}) = \frac{6}{20} = 0.30$

Alternatively, if we designated F_1 and F_2 as the Shakey firms, then we could go to the list of 20 possible outcomes and identify 6 pairs containing just Stable firms: $F_3F_4, F_3F_5, F_4F_3, F_4F_5, F_5F_3, F_5F_4$. Thus, the probability that both firms are Stable is $\frac{6}{20}$

c. $P(\text{One of two firms is Shakey})$
$= P(\text{1st chosen is Shakey and 2nd chosen is Stable})$
$+P(\text{1st chosen is Stable and 2nd chosen is Shakey})$
$= P(\overline{T_1} \cap T_2) + P(T_1 \cap \overline{T_2}) = P(T_2|\overline{T_1})P(\overline{T_1}) + P(\overline{T_2}|T_1)P(T_1)$
$= (\frac{3}{4})(\frac{2}{5}) + (\frac{2}{4})(\frac{3}{5}) = \frac{12}{20} = 0.60$

Alternatively, in the list of 20 outcomes there are 12 pairs which consist of exactly 1 Shakey firm (F_1 or F_2) and exactly 1 Stable firm (F_3 or F_4 or F_5). Thus, the probability of exactly 1 Shakey firm is $\frac{12}{20}$.

d. $P(\text{Both Shakey}) = P(\overline{T_1} \cap \overline{T_2}) = P(\overline{T_2}|\overline{T_1})P(\overline{T_1}) = (\frac{1}{4})(\frac{2}{5}) = \frac{2}{20} = 0.10$

Alternatively, in the list of 20 outcomes there are 2 pairs in which both firms are Shakey (F_1 or F_2). Thus, the probability that both firms are Shakey is $\frac{2}{20}$.

4.14 $P(A) = 192/(192 + 248) = 0.436$; $P(B) = 128/440 = 0.291$;
$P(A \cap B) = 48/440 = 0.109$

4.15 a. $P(A) = P(none \cap high) + P(little \cap high) + P(some \cap high) + P(extensive \cap high)$
$= 0.10 + 0.15 + 0.16 + 0.22 = 0.63$
$P(B) = P(low \cap extensive) + P(medium \cap extensive) + P(high \cap extensive)$
$= 0.04 + 0.10 + 0.22 = 0.36$
$P(C) = P(low \cap none) + P(low \cap little) + P(medium \cap none) + P(medium \cap little)$
$= 0.01 + 0.02 + 0.05 + 0.06 = 0.14$

b. $P(A|B) = P(A \cap B)/P(B) = 0.22/0.36 = 0.611$
$P(A|B) = P(A \cap B)/P(B) = 0.41/(1 - .36) = 0.64$
$P(B|C) = P(B \cap C)/P(C) = 0.14/0.14 = 1$

c. $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.63 + 0.36 - 0.22 = 0.77$
$P(A \cap C) = 0; \quad P(B \cap C) = 0$

4.16 a. P(Both customers pay in full) = (0.70)(0.70) = 0.49

b. P(At least one of 2 customers pay in full) = 1 - P(neither customer pays in full) = 1 - (1-0.70)(1-0.70) = $1 - (0.30)^2 = 0.91$

4.17 Let A = event customer pays 1st month's bill in full and B = event customer pays 2nd month's bill in full. We are given that

$P(A) = 0.70, P(B|A) = 0.95, P(\overline{B}|\overline{A}) = 0.90, P(B|\overline{A}) = 1 - P(\overline{B}|\overline{A}) = 1 - 0.90 = 0.10$

a. $P(A \cap B) = P(B|A)P(A) = (0.95)(0.70) = 0.665$

b. $P(\overline{B} \cap \overline{A}) = P(\overline{B}|\overline{A})P(\overline{A}) = (0.90)(1 - 0.70) = 0.270$

c. P(pay exactly one month in full)
$= 1 - P$(pays neither month or pays both months)
$= 1 - P(\overline{A} \cap \overline{B}) - P(A \cap B) =$ 1 - 0.665 - 0.27 = 0.065

4.18 Let D be the event loan is defaulted, R_1 applicant is poor risk, R_2 fair risk, and R_3 good risk.

$P(D) = 0.01, \quad P(R_1|D) = 0.30, \quad P(R_2|D) = 0.40, \quad P(R_3|D) = 0.30$

$P(\overline{D}) = 0.01, \quad P(R_1|\overline{D}) = 0.10, \quad P(R_2|\overline{D}) = 0.40, \quad P(R_3|\overline{D}) = 0.50$

$P(D|R_1) = \frac{P(R_1|D)P(D)}{P(R_1|D)P(D)+P(R_1|\overline{D})P(\overline{D})} = \frac{(0.30)(0.01)}{(0.30)(0.01)+(0.10)(0.99)} = 0.0294$

4.19 $P(D|R_2) = \frac{P(R_2|D)P(D)}{P(R_2|D)P(D)+P(R_2|\overline{D})P(\overline{D})} = \frac{(0.40)(0.01)}{(0.40)(0.01)+(0.40)(0.99)} = 0.01$

The two probabilities are equal since the proportion of fair risk applicants is the same for both the defaulted and nondefaulted loans.

4.20 Let F be the event fire occurs and T_i be the event a type i furnace is in the home for $i = 1, 2, 3, 4$, where T_4 represent other types.

$P(T_1|F) = \frac{P(F|T_1)P(T_1)}{P(F|T_1)P(T_1)+P(F|T_2)P(T_2)+P(F|T_3)P(T_3)+P(F|T_4)P(T_4)}$
$= \frac{(0.05)(0.30)}{(0.05)(0.30)+(0.03)(0.25)+(0.02)(0.15)+(0.04)(0.30)} = 0.40$

4.21 a. Sensitivity: $P(DA|C) = 50/53 = 0.943$
Specificity: $P(DNA|RO) = 44/47 = 0.936$

b. $P(C|DA) = \frac{(0.943)(0.00108)}{(0.943)(0.00108)+(0.021)(1-0.00108)} = 0.046$

c. $P(RO|DA) = 1 - P(C|DA) = 1 - 0.046 = .954$

d. $P(RO|DNA) = \frac{(0.936)(1-0.00108)}{(0.936)(1-0.00108)+(0.019)(0.00108)} = 0.99998$

4.22 a. $P(y = 3) = \binom{10}{3}(.2)^3(.8)^7 = 0.201$

b. $P(y=2) = \binom{4}{2}(.4)^2(.6)^2 = 0.3456$

c. $P(y=12) = \binom{16}{12}(.7)^{12}(.3)^4 = 0.204$

4.23 a. a. $P(y \leq 4) = P(0) + P(1) + P(2) + P(3) + P(4)$
$= 0.0168 + 0.0896 + 0.2090 + 0.2787 + 0.2322 = 0.8263$

b. $P(y > 4) = 1 - P(y \leq 4) = 1 - 0.8263 = 0.1737$

c. $P(y \leq 7) = 1 - P(8) = 1 - \binom{8}{8}(.4)^8(.6)^{8-8} = 1 - 0.0007 = 0.9993$

d. $P(y > 6) = P(y=7) + P(y=8) = \binom{8}{7}(.4)^7(.6)^{8-7} + \binom{8}{8}(.4)^8(.6)^{8-8} = 0.0085$

4.24 a. Bar graph of P(y)

b. $P(y \leq 2) = P(y=0) + P(y=1) + P(y=2) = 0.5$

c. $P(y \geq 7) = P(y=7) + P(y=8) + P(y=9) + P(y=10) = 0.13$

d. $P(1 \leq y \leq 5) = P(y=1) + P(y=2) + P(y=3) + P(y=4) + P(y=5) = 0.71$

4.25 a. $P(y \geq 3) = 1 - P(y < 3) = 1 - (0.06 + 0.14 + 0.16) = 0.64$

b. $P(2 \leq y \leq 6) = P(y=2) + P(y=3) + P(y=4) + P(y=5) + P(y=6)$
$= 0.16{+}0.14{+}0.12{+}0.10{+}0.08 = 0.60$

c. $P(y > 8) = P(y=9) + P(y=10) = 0.04 + 0.03 = 0.07$

4.26 No, each running of the experiment does not result in one of two possible outcomes.

4.27 No, people may not answer the question.

4.28 Binomial experiment with n=10 and $\pi = 0.60$.

a. $P(y=0) = 0.0001$

b. $P(y=6) = 0.2508$

c. $P(y \geq 6) = 1 - P(y < 6) = 1 - (P(0) + P(1) + P(2) + P(3) + P(4) + P(5))$
$= 1 - (0.3669) = 0.6331$

d. $P(y=10) = 0.0060$

4.29 Binomial experiment with n=10 and $\pi = 0.30$.

a. $P(y=0) = \binom{10}{0}(.3)^0(.7)^{10} = 0.0282$
$P(y=6) = \binom{10}{6}(.3)^6(.7)^4 = 0.0368$
$P(y \geq 6) = 1 - P(y < 6)$
$= 1 - (P(0) + P(1) + P(2) + P(3) + P(4) + P(5)) = 1 - 0.9527 = 0.0473$
$P(y=10) = \binom{10}{10}(.3)^{10}(.7)^0 = 0.000006$

b. With $n = 1000$ and $\pi = .3$ $\quad P(y \leq 100) = \sum_{i=0}^{100} \binom{1000}{i}(.3)^i(.7)^{1000-i}$.
This would be a length calculation. In Section 4.13, we will provide an approximation which will greatly reduce the amount of calculations. However, we can also consider that
$\mu = n\pi = (1000)(.3) = 300, \sigma = \sqrt{n\pi(1-\pi)} = \sqrt{1000(.3)(.7)} = 14.49$
$\mu \pm 3\sigma = 300 \pm (3)(14.49) = (256.53, 343.47)$.
Thus, the chance of observing the event $y \leq 100$ is very small.

4.30 a. $P(y=2) = \binom{10}{2}(.1)^2(.9)^8 = 0.1937$

b. $P(y \geq 2) = 1 - P(y=0) - P(y=1) = 1 - \binom{10}{0}(.1)^0(.9)^{10} - \binom{10}{1}(.1)^1(.9)^9 = 0.2639$

c. $\pi =$ P(either outstanding or good) $= 0.85$; $P(y=8) = \binom{10}{8}(0.85)^8(.15)^2 = 0.2759$

d. $\pi =$ P(unsatisfactory) $= 0.05$; $P(y=0) = \binom{10}{0}(0.05)^0(.95)^{10} = 0.5987$

4.31 No. The trials are not identical.

4.32 Binomial with $n = 50; \pi = .17$

a. $P(y \leq 3) = \sum_{i=0}^{3} \binom{50}{i}(0.07)^i(0.93)^{50-i} = 0.5327$ (using a computer program)

b. The posting of price changes are independent with the same probability 0.07 of being posted incorrectly.

4.33 a. 0.9032 - 0.5000 = 0.4032

b. 0.5000 - 0.0287 = 0.4713

4.34 a. 0.7580 - 0.5000 = 0.2580

b. 0.5000 - 0.1151 = 0.3849

4.35 a. 0.9115 - 0.4168 = 0.4947

b. 0.8849 - 0.6443 = 0.2406

4.36 $z_o = 0$

4.37 $z_o = 1.96$

4.38 $z_o = 2.37$

4.39 $z_o = 1.645$

4.40 $z_o = 1.96$

4.41 a. $P(500 < y < 696) = P(\frac{500-500}{100} < z < \frac{696-500}{100})$
$= P(0 < z < 1.96) = 0.9750 - 0.5000 = 0.475$

b. $P(y > 696) = P(z > \frac{696-500}{100}) = P(z > 1.96) = 1 - 0.9750 = 0.025$

c. $P(304 < y < 696) = P(\frac{304-500}{100} < z < \frac{696-500}{100})$
$= P(-1.96 < z < 1.96) = 0.9750 - 0.0250 = 0.95$

d. $P(500 - k < y < 500 + k) = P(\frac{500-k-500}{100} < z < \frac{500+k-500}{100})$
$= P(-.01k < z < .01k) = 0.60.$
From Table 1 we find,
$P(-.845 < z < .845) = 0.60$. Thus, $.01k = .845 \Rightarrow k = 84.5$

4.42 a. $y < 130 \Rightarrow \frac{y-100}{15} < \frac{130-100}{15} \Rightarrow z < 2$

b. $y > 82.5 \Rightarrow \frac{y-82.5}{15} > \frac{82.5-100}{15} \Rightarrow z > -1.17$

c. $P(y < 130) = P(z < 2) = 0.9772$;
$P(y > 82.5) = P(z > -1.17) = 1 - P(z \leq -1.17) = 0.8790$

d. $P(y > 106) = P(z > \frac{106-100}{15}) = P(z > 0.4) = 1 - P(z \leq 0.4) = 0.3446$
$P(y < 94) = P(z < \frac{94-100}{15}) = P(z < -0.4) = 0.3446$
$P(94 < y < 106) = P(-0.4 < z < 0.4) = 1 - 0.3446 - 0.3446 = 0.3108$

e. $P(y > 130) = P(z > \frac{130-100}{15}) = P(z > 2) = 1 - P(z \leq 2) = 0.0228$
$P(y < 70) = P(z < \frac{70-100}{15}) = P(z < -2) = 0.0228$
$P(70 < y < 130) = P(-2 < z < 2) = 1 - 0.0228 - 0.0228 = 0.9544$

4.43 a. z = 2.33
b. z = 1.28

4.44 a. $P(z > 1.96) = 1 - 0.9750 = 0.025$
b. $P(z > 2.21) = 1 - 0.9864 = 0.0136$
c. $P(z > 2.86) = 1 - 0.9979 = 0.0021$
d. $P(z > 0.73) = 1 = 0.7673 = 0.2327$

4.45 a. $\mu = 39; \sigma = 6; P(y > 50) = P(z > \frac{50-39}{6}) = P(z > 1.83) = 1 - 0.9664 = 0.0336$
b. Since 55 is $\frac{55-39}{6} = 2.67$ std. dev. above $\mu = 39$, thus $P(y > 55) = P(z > 2.67)$ $= 0.0038$. We would then conclude that the voucher has been lost.

4.46 $\mu = 500; \quad \sigma = 100$

a. $P(y > 600) = P(z > \frac{600-500}{100}) = P(z > 1) = 0.1587$
b. $P(y > 700) = P(z > \frac{700-500}{100}) = P(z > 2) = 0.0228$
c. $P(y < 450) = P(z < \frac{450-500}{100}) = P(z < -0.5) = 0.3085$
d. $P(450 < y < 600) = P(\frac{450-500}{100} < z < \frac{600-500}{100}) = P(-0.5 < z < 1) = 0.5328$

4.47 $\mu = 150; \quad \sigma = 35$

a. $P(y > 200) = P(z > \frac{200-150}{35}) = P(z > 1.43) = 0.0764$
b. $P(y > 220) = P(z > \frac{220-150}{35}) = P(z > 2) = 0.0228$
c. $P(y < 120) = P(z < \frac{120-150}{35}) = P(z < -0.86) = 0.1949$
d. $P(100 < y < 200) = P(\frac{100-150}{35} < z < \frac{200-150}{35}) = P(-1.43 < z < 1.43) = 0.8472$

4.48 $\mu = 500; \quad \sigma = 100$

a. Find k such that $P(y > k) = 0.10; \quad P(z > 1.285) = 0.10 \quad \Rightarrow$
$k = 500 + (1.285)(100) = 628.50$ or greater.
b. $P(y < k) = 0.40; \quad P(z < -0.255) = 0.40 \Rightarrow$
$k = 500 + (-0.255)(100) = 474.50$; This is the 40th percentile.

4.49 No. The sample would be biased toward homes for which the homeowner is at home much of the time. For example, the sample would tend to include more people who work at home and retired persons.

4.50 In order to make the sampling random, the network might choose voters based on draws from a random number table, or more simply choose every nth person exiting.

4.51 Start at column 2 line 1 we obtain 150, 729, 611, 584, 255, 465, 143, 127, 323, 225, 483, 368, 213, 270, 062, 399, 695, 540, 330, 110, 069, 409, 539, 015, 564. These would be the women selected for the study.

4.52
a. Use a computer program or the random number table to generate 50 numbers.
b. Reusing the same numbers may create a bias for a certain segment of the population within each of the precincts.

4.53 The sampling distribution would have a mean of 60 and a standard deviation of $\frac{5}{\sqrt{16}} = 1.25$. If the population distribution is somewhat mound-shaped then the sampling distribution of $\bar{y}$ should be approximately mound-shaped. In this situation, we would expect approximately 95% of the possible values of $\bar{y}$ is lie in $60 \pm (2)(1.25) = (57.5, 62.5)$.

4.54 If the population distribution is somewhat mound-shaped then the sampling distribution of $\sum_i y_i$ should be approximately mound-shaped with mean of (60)(16)=960 and a standard deviation of $5\sqrt{16} = 20$. Observing a measurement more than 70 units ($\frac{70}{20} = 3.5$ standard deviations) away from the mean (960) for a normal distribution is very improbable

$(P(|z| > 3.5) \approx 0.0005)$.

4.55 $\mu = 930; \quad \sigma = 130$

a. $P(800 < y < 1100) = P(\frac{800-930}{130} < z < \frac{1100-930}{130}) = P(-1 < z < 1.31)$
$= 0.9049 - 0.1587 = 0.7462$
b. $P(y < 800) = P(z < \frac{800-930}{130}) = 0.1587$
c. $P(y > 1200) = P(z > \frac{1200-930}{130}) = P(z > 2.08) = 1 - 0.9811 = 0.0189$

4.56
a. 125 ± 32 should contain approximately 68% of the weeks
125 ± 64 should contain approximately 95% of the weeks
125 ± 96 should contain approximately 99.7% of the weeks.
b. $P(y > 160) = P(z > \frac{160-125}{32}) = P(z > 1.09) = 1 - 0.8621 = 0.1379$

4.57 $\mu = 2.1; \quad \sigma = 0.3$

a. $P(y > 2.7) = P(z > \frac{2.7-2.1}{0.3}) = P(z > 2) = 0.0228$
b. $P(z > 0.6745) = 0.25 \quad \Rightarrow y_{.75} = 2.1 + (0.6745)(0.3) = 2.30$
c. Let μ_N be the new value of the mean. We need $P(y > 2.7) \leq 0.05$.
From Table 1, $0.05 = P(z > 1.645)$ and $0.05 = P(y > 2.7) = P(\frac{y-\mu_N}{.3} > \frac{2.7-\mu_N}{0.3}) \Rightarrow$
$\frac{2.7-\mu_N}{0.3} = 1.645 \Rightarrow \mu_N = 2.7 - (0.3)(1.645) = 2.2065$

4.58 Mean for fleet is $n\mu = (150)(2.1) = 315$; standard deviation for the fleet is
$\sqrt{n}\sigma = \sqrt{150}(0.3) = 3.67$

4.59 Individual baggage weight has $\mu = 95$; $\sigma = 35$; Total weight has mean $n\mu = (200)(95) = 19,000$; and standard deviation $\sqrt{n}\sigma = \sqrt{200}(35) = 494.97$. Therefore, $P(y > 20,000) = P(z > \frac{20,000-19,000}{494.97}) = P(z > 2.02) = 0.0217$

4.60 No. The last date may not be representative of all days in the month.

4.61 a. $\mu = n\pi = (10000)(0.001) = 10$

b. $\sigma = \sqrt{10000)(0.001)(0.999)} = 3.16$

$P(y < 5) = P(y \le 4) \approx P(z \le \frac{4.5-10}{3.16}) = P(z < -1.74) = 0.0409$

c. $P(y < 2) = P(y \le 1) = P(z \le \frac{1.5-10}{3.16}) = P(z < -2.69) = 0.0036$

4.62 a. $P(y > 2265) = P(z > \frac{2265-2250}{10.2}) = P(z > 1.47) = 0.0707$

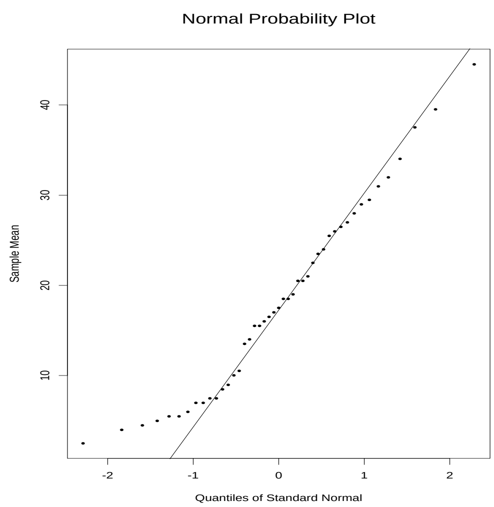
b. Approximately normal with mean $= 2250$ and standard deviation $= \frac{10.2}{\sqrt{15}} = 2.63$

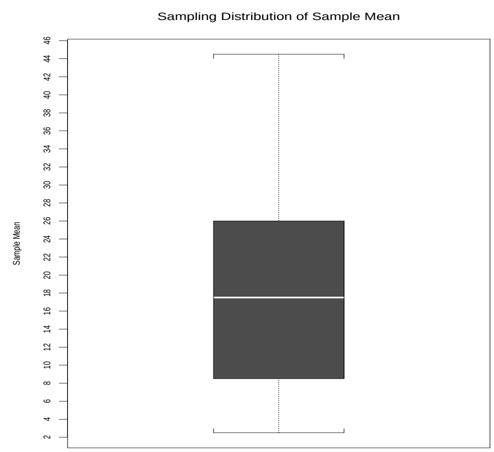
4.63 $\sigma_{\bar{y}} = \frac{10.2}{\sqrt{15}} = 2.63$

4.64 $P(\bar{y} \ge 2268) = P(z \ge \frac{2268-2250}{2.63}) = P(z \ge 6.84) \approx 0.$

4.65 No, there is strong evidence that the new fabric has a greater mean breaking strength.

4.66 a. The normal probability plot and box plots are given here.





Note that the plotted points deviate from the straight-line.

b. Since the population distribution is skewed to the right for this sample and with the sample size only 2, the sampling distribution for $\bar{y}$ will be skewed to the right.

4.67 Could use random sampling by first identifying all returns with income greater than \$15,000. Next, create a computer file with all such returns listed in alphabetical order by last name of client. Randomly select 1% of these returns using a systematic random sampling technique.

4.68 a. The random variable y, representing the number of sales in 5 calls, is binomial with $\pi = 0.01$. The event that the first sale occurs in the first five calls is the complement of the event that no sales occur in the first five calls. Thus, the probability of interest is $1 - P(y = 0) = 1 - \binom{5}{0}(.01)^0(.99)^5 = 0.049$

b. If the first sale occurs after 10 calls, the first ten calls must have resulted in no sales. Let y be the number of sales in the 10 calls, $P(y = 0) = \binom{10}{0}(.01)^0(.99)^{10} = 0.9044$

4.69 $n = 400$ $\pi = 0.2$

a. $\mu = n\pi = 400(.20) = 80; \sigma = \sqrt{400(.2)(.8)} = 8$
$P(y \leq 25) \approx P(z \leq \frac{25-80}{8}) = P(z \leq -6.875) \approx 0.$

b. The ad is not successful. With $\pi = .20$, we expect 80 positive responses out of 400 but we observed only 25. The probability of getting so few positive responses is virtually 0 if $\pi = .20$. We therefore conclude that π is much less than 0.20.

4.70 $n = 20000, \pi = 0.0001$. There are 2 possible outcomes and each birth is an independent event. We cannot use the normal approximation because $n\pi = (20000)(0.0001) = 2 < 5$. We can use the binomial formula:

$$P(y \geq 1) = 1 - P(y = 0) = 1 - \binom{20000}{0}(.0001)^0(.9999)^{20000} = .8647$$

4.71 a.,b. The mean and standard deviation of the sampling distribution of $\bar{y}$ are given when the population distribution has values $\mu = 100, \sigma = 15$:

Sample Size	Mean	Standard Deviation
5	100	6.708
20	100	3.354
80	100	1.677

c. As the sample size increases, the sampling distribution of $\bar{y}$ concentrates about the true value of μ. For $n = 5$ and 20, the values of $\bar{y}$ could be a considerable distance from 100.

4.72 a.-c. The probabilites are given here:

Sample Size	$P(\bar{y} \geq 105)$	$P(\bar{y} \leq 95)$	$P(95 \leq \bar{y} \leq 105)$
5	0.2280	0.2280	0.5439
20	0.0680	0.0680	0.8640
80	0.0014	0.0014	0.9971

The sample mean's chance of being close to the population mean increases very rapidly as the sample size increases. Of course, the rate of increase also depends on the population standard deviation, σ.