Answers

1. Here are the calculations:

$$E(D) = (2.5)(.19) + (7.5)(.29) + (12.5)(.34) + (20)(.18) = 10.5$$
$$SD(D) = \sqrt{(2.5-10.5)^2(.19) + (7.5-10.5)^2(.29) + (12.5-10.5)^2(.34) + (20-10.5)^2(.18)} = 5.69$$

Whether this is reasonable or not depends on how you look at it. Over the past 35 years or so, the stock market has returned roughly 8% per year, so considering 10.5% as being enough for "good" seems a bit greedy. On the other hand, over the last 20 years or so, the market has returned 13–15% annually, so maybe 10.5% isn't so unreasonable. Note the relatively high standard deviation, which comes from the 37% of the people who are happy with relatively little (2.5%) or quite a bit (20%).

2.(a) Let S be the repair score of a book. Then

$$E(S) = (1)(.815) + (2)(.044) + (3)(.107) + (4)(.034) = 1.36.$$

(b) Let C be the repair cost of a book. Then

$$E(C) = (0)(.815) + (25)(.044) + (75)(.107) + (500)(.034) = \$26.125 \text{ per book}$$

In fact, these costs are understated, since they do not reflect many of the labor costs involved in book repair, and especially that extremely rare volumes are far more expensive to repair.

3. Note that $D \sim B(n = 7, p = .287)$.
(a) $E(D) = np = (7)(.287) = 2.009$
$SD(D) = \sqrt{np(1-p)} = \sqrt{(7)(.287)(.713)} = 1.197$
(b) The Binomial probability function has the form

$$P(D = k) = \binom{n}{k} p^k (1-p)^{n-k}, \qquad k = 0, 1, 2, \ldots, n.$$

Here $n = 7$, $p = .287$ and $k = 2$, so $P(D = 2) = .31874$.
(c) Now we want

$$P(D \geq 3) = 1-P(D < 3) = 1-[P(D = 0)+P(D = 1)+P(D = 2)] = 1-(.09368+.26395+.31874) = .32363.$$

(d) Let G be the number of research groups with at least three of the people being Ph.D.'s. Then, by part (c), $G \sim B(n = 4, p = .32363)$. We want $P(G = 2) = \binom{4}{2}(.32363)^2(.67637)^2 = .28749$.

4. Let R be the daily S&P return; we're told that $R \sim N(.00032, .00859^2)$.
(a)

$$P(R > .01) = P\left(Z > \frac{.01 - .00032}{.00859}\right) = P(Z > 1.13) = .1292$$

(b)

$$P(R < -.01) = P\left(Z < \frac{-.01 - .00032}{.00859}\right) = P(Z < -1.20) = .1151$$

(c)

$$P(R > 0) = P\left(Z > \frac{0 - .00032}{.00859}\right) = P(Z > -.04) = .5160$$

(d)

$$P(-.015 < R < .015) = P\left(\frac{-.015 - .00032}{.00859} < Z < \frac{.015 - .00032}{.00859}\right) = P(-1.78 < Z < 1.71) = .9189$$

(e)

$$P(R < .025) = P\left(Z < \frac{.025 - .00032}{.00859}\right) = P(Z < 2.87) = .9979$$

(f) Let M be the daily return of the mutual fund. We're told that $M \sim N(.00032, \sigma^2)$, and that $P(M > 0) = .7881$. Thus

$$\begin{aligned} .7881 &= P(M > 0) \\ &= P\left(Z > \frac{0 - .00032}{\sigma}\right) \end{aligned}$$

Thus, $-.00032/\sigma = z_{.7881} = -z_{.2119} = -.80$, which means that $\sigma = .00032/.80 = .0004$. This is more than 20 times smaller than the market standard deviation, so the claim seems highly dubious.

5.(a) Since the area of a rectangle is width $\times$ height, the area of the entire map is $m \times m = m^2$. Each of the non–Murphy zones is a rectangle with width $m/2 - 2b$ and height $m - 2b$, so the total area of the non–Murphy zones is

$$\begin{aligned} 2(m/2 - 2b)(m - 2b) &= (m - 4b)(m - 2b) \\ &= m^2 - 6mb + 8b^2 \\ &= m^2[1 - 6(b/m) + 8(b/m)^2] \end{aligned}$$

(b) From (a) the area of the Murphy zone is $m^2 - m^2[1 - 6(b/m) + 8(b/m)^2]$, so the probability of a randomly chosen location falling in the Murphy zone is $6(b/m) - 8(b/m)^2$.

(c) Substituting $b/m = .1$ into the formula yields .52 for the probability; that is, a point picked at random has a better than even chance of ending up in a Murphy zone of width just one–tenth that of the entire map. This certainly seems to be a surprising result. The reason is that the Murphy zone tracks the outermost, and hence largest, dimensions of the map, so a relatively narrow strip still encloses a comparatively large total area. That is, it isn't just your impression that locations tend to be along the edges of a map; they really do! Other examples of Murphy's Law have scientific explanations as well; for example, buttered toast that falls off a table really **is** more likely to fall buttered–side down. See "Murphy's Law of Maps," by Robert A.J. Matthews, *Teaching Statistics*, **19**, 34–35, for a discussion of this example.

(d) D, the number of destinations in the Murphy zone, is binomially distributed with $n = 40$ and $p = .52$.

(e) $E(D) = np = (40)(.52) = 20.8$

(f) $SD(D) = \sqrt{np(1-p)} = \sqrt{(40)(.52)(.48)} = 3.16$

(g) We use the normal approximation to the binomial for this:

$$\begin{aligned} P(D \geq 15) &= P(D > 14.5) \\ &\approx P\left(Z > \frac{14.5 - 20.8}{3.16}\right) \\ &= P(Z > -1.994) \\ &= .9769 \end{aligned}$$

(h)

$$\begin{aligned} P(D = 19) &= P(18.5 < D < 19.5) \\ &\approx P\left(\frac{18.5 - 20.8}{3.16} < Z < \frac{19.5 - 20.8}{3.16}\right) \\ &= P(-.728 < Z < -.411) \\ &= .1072 \end{aligned}$$

My scientific calculator gives an exact value of .1067 for this value, which is less than 1/2 of one per cent different from the approximate value.

6. Let L be the load of a "top" mutual fund. Then

$$E(L) = (0)(.34) + (3)(.3) + (4.75)(.08) + (5.5)(.1) + (5.75)(.1) + (7)(.08) = 2.965$$

$$SD(L) = \sqrt{(0^2)(.34) + (3^2)(.3) + (4.75^2)(.08) + (5.5^2)(.1) + (5.75^2)(.1) + (7^2)(.08) - 2.965^2} = 2.44$$

Given the observed distribution, either of the modal values (0 or 3) would seem to be the best choice for a "typical value." In fact, I would probably resist the urge to give one number, and instead reply "There are two possibilities: a no load fund (about one–third of the funds) or a load fund; if it's the latter, more funds have a 3% load than any other value."

7. Examination of the two distributions shows that both have a modal value at "Fairly close," so in that sense they're similar. The extreme levels ("Very close" and "Very distant") are higher probability events for women than men, though, so apparently men are more moderate in their feelings than women. If we assign scores $F = \{-3, -1, 1, 3\}$ to the four answers (these are symmetric around zero) numerical summaries confirm these patterns, as the means are similar and negative (on the side of closeness), while the standard deviation for women is noticeably larger than that for men:

Women

$$E(F) = (-3)(.26) + (-1)(.33) + (1)(.25) + (3)(.16) = -0.38$$

$$SD(F) = \sqrt{(-3)^2(.26) + (-1)^2(.33) + (1)^2(.25) + (3)^2(.16) - (-0.38)^2} = 2.05$$

Men

$$E(F) = (-3)(.23) + (-1)(.38) + (1)(.28) + (3)(.11) = -0.46$$

$$SD(F) = \sqrt{(-3)^2(.23) + (-1)^2(.38) + (1)^2(.28) + (3)^2(.11) - (-0.46)^2} = 1.87$$

Another way to quantify this is using the coefficient of variation: it is 5.39 for women and 4.06 for men, supporting the conclusion that the distribution for women is more variable.

8.(a) $W \sim B(5, .025)$, so $E(W) = (5)(.025) = .125$, while $SD(W) = \sqrt{(5)(.025)(.975)} = .349$.

(b)

$$P(W = 2) = \binom{5}{2}(.025)^2(.975)^3 = .0058$$

(c)

$$P(W = 0) = \binom{5}{0}(.025)^0(.975)^5 = .881$$

(d) Let C be the number of companies with no women among the top five earners. Then $C \sim B(10, .881)$ (based on the answer to part (c)). Then,

$$P(C = 10) = \binom{10}{10}(.881)^{10}(.119)^0 = .282$$

9. Let R be the daily S&P return; we're told that $R \sim N(.00039, .00721^2)$.

(a)

$$P(R < .01) = P\left(Z < \frac{.01 - .00039}{.00721}\right) = P(Z < 1.33) = .9082$$

(b)

$$P(R > 0) = P\left(Z > \frac{0 - .00039}{.00721}\right) = P(Z > -.05) = .5199$$

(c)

$$P(-.01 < R < .01) = P\left(\frac{-.01 - .00039}{.00721} < Z < \frac{.01 - .00039}{.00721}\right) = P(-1.44 < Z < 1.33) = .8333$$

(d)

$$P(R < .005) = P\left(Z < \frac{.005 - .00039}{.00721}\right) = P(Z < .64) = .7389$$

(f) Let M be the daily return of the mutual fund. We're told that $M \sim N(\mu, .00721^2)$, and that $P(M > 0) = .7881$. Thus

$$\begin{aligned} .7881 &= P(M > 0) \\ &= P\left(Z > \frac{0 - \mu}{.00721}\right) \end{aligned}$$

Thus, $-\mu/.00721 = z_{.7881} = -z_{.2119} = -.80$, which means that $\mu = (.80)(.00721) = .00577$. This is almost 15 times larger than the market average, so the claim seems highly dubious (it would translate to an annual return of 265%! Sign me up!).

10. Here are the calculations:

$$\begin{aligned} E(M) &= (2)(.0009) + (1)(.0278) + (0)(.8442) + (-1)(.0937) + (-2)(.0046) = -.0733 \\ SD(M) &= \{[2 - (-.0733)]^2(.0009) + [1 - (-.0733)]^2(.0278) + [0 - (-.0733)]^2(.8442) \\ &\quad + [-1 - (-.0733)]^2(.0937) + [-2 - (-.0733)]^2(.0046)\}^{.5} = .375 \end{aligned}$$

Thus, A bonds moved down slightly on average, with standard deviation much higher than the mean movement. If I were to give a "typical" value, however, I would report the mode — almost 85% of A bonds were still A bonds one year later.

11. Note that $B \sim Bin(n = 6, p = .41)$.
(a) $E(B) = np = (6)(.41) = 2.46$
$SD(B) = \sqrt{np(1-p)} = \sqrt{(6)(.41)(.59)} = 1.205$
(b) The Binomial probability function has the form

$$P(B = k) = \binom{n}{k} p^k (1-p)^{n-k}, \qquad k = 0, 1, 2, \ldots, n.$$

Here $n = 6$, $p = .41$ and $k = 2$, so $P(B = 2) = .3055$.
(c) Now we want $P(B > 0) = 1 - P(B = 0) = 1 - .59^6 = 1 - .0422 = .9578$.
(d) The number of binge drinkers B is now Binomially distributed with $n = 500$ and $p = .41$. We'll approximate the probability we want using the normal approximation to the Binomial, making sure not to forget to use the continuity correction. We know that $E(B) = (500)(.41) = 205$ and $SD(B) = \sqrt{(500)(.41)(.59)} = 11.0$.

$$\begin{aligned} P(B > 230) &= P(B \geq 230.5) \\ &\approx P\left(Z \geq \frac{230.5 - 205}{11}\right) \\ &= P(Z \geq 2.32) = .0102. \end{aligned}$$

12. Let R be the monthly S&P return; we're told that $R \sim N(.01784, .04793^2)$.
(a)

$$P(R > .01) = P\left(Z > \frac{.01 - .01784}{.04793}\right) = P(Z > -.16) = .5636$$

(b)

$$P(R < 0) = P\left(Z < \frac{0 - .01784}{.04793}\right) = P(Z < -.37) = .3557$$

(c)

$$P(-.05 < R < .05) = P\left(\frac{-.05 - .01784}{.04793} < Z < \frac{.05 - .01784}{.04793}\right) = P(-1.42 < Z < .67) = .6708$$

(d) Let M be the monthly return of the mutual fund. We're told that $M \sim N(.01784, \sigma^2)$, and that $P(M > 0) = .9452$. Thus

$$\begin{aligned} .9452 &= P(M > 0) \\ &= P\left(Z > \frac{0 - .01784}{\sigma}\right) \end{aligned}$$

Thus, $-.01784/\sigma = z_{.0548} = -z_{.9452} = -1.6$, which means that $\sigma = .01784/1.6 = .01115$. This is more than 4 times smaller than the market standard deviation, so the claim seems dubious.