Chapter 4: Special Distributions

Section 4.2: The Poisson Distribution

4.2.1 $p = P(\text{word is misspelled}) = \dfrac{1}{3250}$; $n = 6000$. Let x = number of words misspelled. Using the exact binomial analysis, $P(X = 0) = \dbinom{6000}{0}\left(\dfrac{1}{3250}\right)^0\left(\dfrac{3249}{3250}\right)^{6000} = 0.158$. For the Poisson approximation, $\lambda = 6000\left(\dfrac{1}{3250}\right) = 1.846$, so $P(X = 0) \doteq \dfrac{e^{-1.846}(1.846)^0}{0!} = 0.158$. The agreement is not surprising because n is so large and p is so small (recall Example 4.2.1).

4.2.2 Let X = number of prescription errors. Then $\lambda = np = 10 \cdot \dfrac{905}{289{,}411} = 0.0313$, and $P(X \geq 1) = 1 - P(X = 0) \doteq 1 - \dfrac{e^{-0.0313}(0.0313)^0}{0!} = 0.031$.

4.2.3 Let X = number born on Poisson's birthday. Since $n = 500$, $p = \dfrac{1}{365}$, and $\lambda = 500 \cdot \dfrac{1}{365} = 1.370$, $P(X \leq 1) = P(X = 0) + P(X = 1) \doteq \dfrac{e^{-1.370}(1.370)^0}{0!} + \dfrac{e^{-1.370}(1.370)^1}{1!} = 0.602$.

4.2.4 **(a)** Let X = number of chromosome mutations. Given that n = 20,000 and $p = \dfrac{1}{10{,}000}$ (so $\lambda = 2$), $P(X = 3) \doteq e^{-2}2^3/3! = 0.18$.

(b) Listed in the table are values of $P(X \geq k)$ under the assumption that $p = \dfrac{1}{10{,}000}$. If X is on the order of 5 or 6, the credibility of that assumption becomes highly questionable.

k	$P(X \geq k)$
3	0.3233
4	0.1429
5	0.0527
6	0.0166

4.2.5 Let X = number of items requiring a price check. If $p = P(\text{item requires price check}) = 0.01$ and $n = 10$, a binomial analysis gives $P(X \geq 1) = 1 - P(X = 0) = 1 - \dbinom{10}{0}(0.01)^0(0.99)^{10} = 0.10$. Using the Poisson approximation, $\lambda = 10(0.01) = 0.1$ and $P(X \geq 1) = 1 - P(X = 0)$ $\doteq 1 - \dfrac{e^{-0.1}(0.1)^0}{0!} = 0.10$. The exact model that applies here is the hypergeometric, rather than the binomial, because p is a function of the previous items purchased. However, the variation in p will be essentially zero for the 10 items purchased, so the binomial and hypergeometric models in this case will be effectively the same.

4.2.6 Let X = number of policy-holders who will die next year. Since $n = 120$, $p = \frac{1}{150}$, and $\lambda = \frac{120}{150} = 0.8$, P(company will pay at least \$150,000 in benefits) = $P(X \geq 3)$ = $1 - P(X \leq 2) \doteq 1 - \sum_{k=0}^{2} \frac{e^{-0.8}(0.8)^k}{k!} = 0.047$.

4.2.7 Let X = number of pieces of luggage lost. Given that $n = 120$, $p = \frac{1}{200}$, (so $\lambda = 120 \cdot \frac{1}{200} = 0.6$), $= P(X \geq 2) = 1 - P(X \leq 1) = 1 - \sum_{k=0}^{1} \frac{e^{-0.6}(0.6)^k}{k!} = 0.122$.

4.2.8 Let X = number of cancer cases. If $n = 9500$ and $p = \frac{1}{1{,}000{,}000}$, then $\lambda = \frac{9{,}500}{1{,}000{,}000} = 0.0095$ and $P(X \geq 2) = 1 - P(X \leq 1) = 1 - \sum_{k=0}^{1} \frac{e^{-0.0095}(0.0095)^k}{k!} = 0.00005$. The fact that the latter is so small suggests that a lineman's probability of contracting cancer is considerably higher than the one in a million value for p characteristic of the general population.

4.2.9 Let X = number of solar systems with intelligent life and let $p = P$(solar system is inhabited). For n = 100,000,000,000, $P(X \geq 1) = 1 - P(X = 0) = 1 - \binom{100{,}000{,}000{,}000}{0} p^0 \cdot (1-p)^{100{,}000{,}000{,}000}$. Solving $1 - (1-p)^{100{,}000{,}000{,}000} = 0.50$ gives $p = 6.9 \times 10^{-12}$. Alternatively, it must be true that $1 - \frac{e^{-\lambda}\lambda^0}{0!} = 0.50$, which implies that $\lambda = -\ln(0.50) = 0.69$. But $0.69 = np = 1 \times 10^{11} \cdot p$, so $p = 6.9 \times 10^{-12}$.

4.2.10 The average number of fatalities per corps-year = $\frac{109(0) + 65(1) + 22(2) + 3(3) + 1(4)}{200} = 0.61$, so the presumed Poisson model is $p_X(k) = \frac{e^{-0.61}(0.61)^k}{k!}$, $k = 0, 1, \ldots$ Evaluating $p_X(k)$ for $k = 0, 1, 2, 3$, and 4+ shows excellent agreement between the observed proportions and the corresponding Poisson probabilities.

No. of deaths, k	Frequency	Proportion	$p_X(k)$
0	109	0.545	0.5434
1	65	0.325	0.3314
2	22	0.110	0.1011
3	3	0.015	0.0206
4+	1	0.005	0.0035
	200	1.000	1.0000

4.2.11 The observed number of major changes = 0.44 (= $\bar{x} = \frac{1}{356}$[237(0) + 90(1) + 22(2) + 7(3)]), so the presumed Poisson model is $p_X(k) = \frac{e^{-0.44}(0.44)^k}{k!}$, $k = 0, 1, \ldots$ Judging from the agreement evident in the accompanying table between the set of observed proportions and the values for $p_X(k)$, the hypothesis that X is a Poisson random variable is entirely credible.

No. of changes, k	Frequency	Proportion	$p_X(k)$
0	237	0.666	0.6440
1	90	0.253	0.2834
2	22	0.062	0.0623
3+	7	0.020	0.0102
	356	1.000	1.0000

4.2.12 Since $\bar{x} = \frac{1}{40}[9(0)+13(1)+10(2)+5(3)+2(4)+1(5)] = 1.53$, $p_X(k) = \frac{e^{-1.53}(1.53)^k}{k!}$, $k = 0, 1, \ldots$ Yes, the Poisson appears to be an adequate model, as indicated by the close agreement between the observed proportions and the values of $p_X(k)$.

No. of bags lost, k	Frequency	Proportion	$p_X(k)$
0	9	0.225	0.2165
1	13	0.325	0.3313
2	10	0.250	0.2534
3	5	0.125	0.1293
4	2	0.050	0.0494
5+	1	0.025	0.0201
	40	1.000	1.0000

4.2.13 The average of the data is $\frac{1}{113}[82(0)+25(1)+4(2)+0(3)+2(4) \doteq 0.363$. Then use the model $e^{-0.363}\frac{0.363^k}{k!}$. Usual statistical practice suggests collapsing the low frequency categories, in this case, $k = 2, 3, 4$. The result is the following table.

No. of countires, k	Frequency	$p_X(k)$	Expected frequency
0	82	0.696	78.6
1	25	0.252	28.5
2+	6	0.052	5.9

The level of agreement between the observed and expected frequencies suggests that the Poisson is a good model for these data.

4.2.14 **(a)** The model $p_X(k) = e^{-2.157}(2.157)^k/k!$, $k = 0, 1, \ldots$ fits the data fairly well (where $\bar{x} = 2.157$), but there does appear to be a slight tendency for deaths to "cluster"— that is, the values 0, 5, 6, 7, 8, and 9 are all over-represented.

No. of deaths, k	Frequency	$p_X(k)$	Expected frequency
0	162	0.1157	126.8
1	267	0.2495	273.5
2	271	0.2691	294.9
3	185	0.1935	212.1
4	111	0.1043	114.3
5	61	0.0450	49.3
6	27	0.0162	17.8
7	8	0.0050	5.5
8	3	0.0013	1.4
9	1	0.0003	0.3
10+	0	0.0001	0.1

(b) Deaths may not be independent events in all cases, and the fatality rate may not be constant.

4.2.15 If the mites exhibit any sort of "contagion" effect, the independence assumption implicit in the Poisson model will be violated. Here, $\bar{x} = \frac{1}{100}[55(0) + 20(1) + \ldots + 1(7)] = 0.81$, but $p_X(k) = e^{-0.81}(0.81)^k/k!$, $k = 0, 1, \ldots$ does not adequately approximate the infestation distribution.

No. of infestations, k	Frequency	Proportion	$p_X(k)$
0	55	0.55	0.4449
1	20	0.20	0.3603
2	21	0.21	0.1459
3	1	0.01	0.0394
4	1	0.01	0.0080
5	1	0.01	0.0013
6	0	0	0.0002
7+	1	0.01	0.0000
		1.00	1.0000

4.2.16 Let X = number of repairs needed during an eight-hour workday. Since $E(X) = \lambda = 8 \cdot \frac{1}{5} = 1.6$, $P(\text{expenses} \leq \$100) = P(X \leq 2) = \sum_{k=0}^{2} \frac{e^{-1.6}(1.6)^k}{k!} = 0.783$.

4.2.17 Let X = number of transmission errors made in next half-minute. Since $E(X) = \lambda = 4.5$, $P(X > 2) = 1 - P(X \leq 2) = 1 - \sum_{k=0}^{2} \frac{e^{-4.5}(4.5)^k}{k!} = 0.826$.

4.2.18 If $P(X = 0) = e^{-\lambda}\lambda^0/0! = e^{-\lambda} = \frac{1}{3}$, then $\lambda = 1.10$. Therefore, $P(X \geq 2) = 1 - P(X \leq 1) = 1 - e^{-1.10}(1.10)^0/0! - e^{-1.10}(1.10)^1/1! = 0.301$.

4.2.19 Let X = number of flaws in 40 sq. ft. Then $E(X) = 4$ and $P(X \geq 3) = 1 - P(X \leq 2) =$
$1 - \sum_{k=0}^{2} \frac{e^{-4}4^k}{k!} = 0.762$.

4.2.20 Let X = number of particles counted in next two minutes. Since the rate at which the particles are counted <u>per minute</u> is $4.017\left(= \frac{482}{120}\right)$, $E(X) = 8.034$ and $P(X = 3) = \frac{e^{-8.034}(8.034)^3}{3!} = 0.028$.
Now, suppose X = number of particles counted <u>in one minute</u>. Then P(3 particles are counted in next two minutes) $= P(X = 3) \cdot P(X = 0) + P(X = 2) \cdot P(X = 1) + P(X = 1) \cdot P(X = 2) + P(X = 0) \cdot P(X = 3) = 0.028$, where $\lambda = 4.017$.

4.2.21 **(a)** Let X = number of accidents in next five days. Then $E(X) = 0.5$ and $P(X = 2) = e^{-0.5}(0.5)^2/2! = 0.076$.

(b) No. P(4 accidents occur during next two weeks) =
$P(X = 4) \cdot P(X = 0) + P(X = 3) \cdot P(X = 1) + P(X = 2) \cdot P(X = 2) + P(X = 1) \cdot P(X = 3) + P(X = 0) \cdot P(X = 4)$.

4.2.22 If $P(X = 1) = P(X = 2)$, then $e^{-\lambda}\lambda^1/1! = e^{-\lambda}\lambda^2/2!$, which implies that $2\lambda = \lambda^2$, or, equivalently, $\lambda = 2$. Therefore, $P(X = 4) = e^{-2}2^4/4! = 0.09$.

4.2.23 $P(X \text{ is even}) = \sum_{k=0}^{\infty} \frac{e^{-\lambda}\lambda^{2k}}{(2k)!} = e^{-\lambda}\left\{1 + \frac{\lambda^2}{2!} + \frac{\lambda^4}{4!} + \frac{\lambda^6}{6!} + \cdots\right\} = e^{-\lambda} \cdot \cosh \lambda =$
$e^{-\lambda}\left(\frac{e^{\lambda} + e^{-\lambda}}{2}\right) = \frac{1}{2}(1 + e^{-2\lambda})$.

4.2.24 $f_{X+Y}(w) = \sum_{k=0}^{\infty} p_k(k)p_Y(w-k) = \sum_{k=0}^{w} e^{-\lambda}\frac{\lambda^k}{k!}e^{-\mu}\frac{\mu^{w-k}}{(w-k)!}$
$= e^{-(\lambda+\mu)}\sum_{k=0}^{w}\frac{1}{k!(w-k)!}\lambda^k\mu^{w-k} = e^{-(\lambda+\mu)}\frac{1}{w!}\sum_{k=0}^{w}\frac{w!}{k!(w-k)!}\lambda^k\mu^{w-k} = e^{-(\lambda+\mu)}\frac{1}{w!}(\lambda+\mu)^w$
The last expression is the w^{th} term of the Poisson pdf with parameter $\lambda + \mu$.

4.2.25 From Definition 3.11.1 and Theorem 3.7.1, $P(X_2 = k) = \sum_{x_1=k}^{\infty}\binom{x_1}{k}p^k(1-p)^{x_1-k} \cdot \frac{e^{-\lambda}\lambda^{x_1}}{x_1!}$.
Let $y = x_1 - k$. Then $P(X_2 = k) = \sum_{y=0}^{\infty}\binom{y+k}{k}p^k(1-p)^y \cdot \frac{e^{-\lambda}\lambda^{y+k}}{(y+k)!} = \frac{e^{-\lambda}(\lambda p)^k}{k!} \cdot \sum_{y=0}^{\infty}\frac{[\lambda(1-p)]^y}{y!} =$
$\frac{e^{-\lambda}(\lambda p)^k}{k!} \cdot e^{\lambda(1-p)} = \frac{e^{-\lambda p}(\lambda p)^k}{k!}$.

4.2.26 **(a)** Yes, because the Poisson assumptions are probably satisfied—crashes are independent events and the crash rate is likely to remain constant.

(b) Since $\lambda = 2.5$ crashes per year, $P(X \geq 4) = 1 - P(X \leq 3) = 1 - \sum_{k=0}^{3}\frac{e^{-2.5}(2.5)^k}{k!} = 0.24$.

(c) Let Y = interval (in yrs.) between next two crashes. By Theorem 4.2.3,
$P(Y < 0.25) = \int_0^{0.25} 2.5e^{-2.5y}dy = 1 - 0.535 = 0.465.$

4.2.27 Let X = number of deaths in a week. Based on the daily death rate, $E(X) = \lambda = 0.7$. Let Y = interval (in weeks) between consecutive deaths. Then $P(Y > 1) = \int_1^{\infty} 0.7e^{-0.7y}dy = -e^{-u}\Big|_{0.7}^{\infty}$ $= 0.50$.

4.2.28 Given that $f_Y(y) = 0.027e^{-0.027}$, $P(Y_1 + Y_2 < 40) = \int_0^{40}(0.027)^2 ye^{-0.027y}dy = \int_0^{1.08} ue^{-u}du =$
$e^{-u}(-u-1)\Big|_0^{1.08} = 1 - 0.706 = 0.29$ (where $u = 0.027y$).

4.2.29 Let X = number of bulbs burning out in 1 hour. Then $E(X) = \lambda = 0.11$. Let Y = number of hours a bulb remains lit. Then $P(Y < 75) = \int_0^{75} 0.011e^{-0.011y}dy = -e^{-u}\Big|_0^{0.825} = 0.56$. (where $u = 0.011y$).
Since $n = 50$ bulbs are initially online, the expected number that will fail to last at least 75 hours is $50 \cdot P(Y < 75)$, or 28.

4.2.30 Assume that 29 long separations and 7 short separations are to be randomly arranged. In order for "bad things to come in fours" three of the short separations would have to occur at least once in the 30 spaces between and around the 29 long separations. For that to happen, either (1) 3 short separations have to occur in one space and the remaining 4 shorts in 4 other spaces (2) 3 short separations occur in one space, 2 short separations occur in another space, and 1 short separation occurs in each of two spaces or (3) 3 short separations occur in each of two spaces and the remaining short occurs in a third space. The combined probability of these three possibilities is

$$\frac{\binom{30}{5}\binom{5}{1} + \binom{30}{4}\frac{4!}{2!1!1!} + \binom{30}{3}\binom{3}{1}}{\binom{36}{29}} = 0.126$$

Section 4.3: The Normal Distribution

4.3.1 **(a)** 0.5782 **(b)** 0.8264 **(c)** 0.9306 **(d)** 0.0000

4.3.2 **(a)** $0.9808 - 0.5000 = 0.4808$
(b) $0.4562 - 0.2611 = 0.1951$
(c) $1 - 0.1446 = 0.8554 = P(Y < 1.06)$
(d) 0.0099
(e) $P(Z \geq 4.61) < P(Z \geq 3.9) = 1 - 1.0000 = 0.0000$

4.3.3 **(a)** Both are the same because of the symmetry of $f_Z(z)$.

(b) Since $f_Z(z)$ is decreasing for all $z > 0$, $\int_{a-\frac{1}{2}}^{a+\frac{1}{2}} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz$ is larger than $\int_{a}^{a+1} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz$.

4.3.4 **(a)** $\int_0^{1.24} e^{-z^2/2} dz = \sqrt{2\pi} \int_0^{1.24} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = \sqrt{2\pi}(0.8925 - 0.5000) = 1.234$

(b) $\int_{-\infty}^{\infty} 6e^{-z^2/2} dz = 6\sqrt{2\pi} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = 6\sqrt{2\pi}$

4.3.5 **(a)** –0.44 **(b)** 0.76 **(c)** 0.41 **(d)** 1.28 **(e)** 0.95

4.3.6 From Appendix Table A.1, $z_{.25} = 0.67$ and $z_{.75} = -0.67$, so $Q = 0.67 - (-0.67) = 1.34$.

4.3.7 Let X = number of decals purchased in November. Then X is binomial with n = 74,806 and $p = 1/12$.
$P(50X < 306{,}000) = P(X < 6120) = P(X \le 6119)$. Using the DeMoivre-Laplace approximation with continuity correction gives

$$P(X \le 6119) \doteq P\left(Z \le \frac{6119.5 - 74{,}806(1/12)}{\sqrt{74{,}806(1/12)(11/12)}}\right) = P(Z \le -1.51) = 0.0655$$

4.3.8 Let X = number of usable cabinets in next shipment. Since $np = 1600(0.80) = 1280$ and $\sqrt{np(1-p)} = \sqrt{1600(0.80)(0.20)} = 16$, P(shipment causes no problems) = $P(1260 \le X \le 1310)$ =

$$P\left(\frac{1259.5-1280}{16} \le \frac{X-1280}{16} \le \frac{1310.5-1280}{16}\right) \doteq P(-1.28 \le Z \le 1.91) = 0.8716.$$

4.3.9 Let X = number of voters challenger receives. Given that $n = 400$ and $p = P$(voter favors challenger) = 0.45, $np = 180$ and $np(1-p) = 99$.

(a) $P(\text{tie}) = P(X = 200) = P(199.5 \le X \le 200.5) =$

$$P\left(\frac{199.5-180}{\sqrt{99}} \le \frac{X-180}{\sqrt{99}} \le \frac{200.5-180}{\sqrt{99}}\right) \doteq P(1.96 \le Z \le 2.06) = 0.0053.$$

(b) $P(\text{challenger wins}) = P(X > 200) = P(X \ge 200.5) =$

$$P\left(\frac{X-180}{\sqrt{99}} \ge \frac{200.5-180}{\sqrt{99}}\right) \doteq P(Z \ge 2.06) = 0.0197.$$

4.3.10 **(a)** Let X = number of shots made in next 100 attempts.
Since $p = P$(attempt is successful) = 0.70, $P(75 \le X \le 80)$ =

$$\sum_{k=75}^{80} \binom{100}{k} (0.70)^k (0.30)^{100-k}.$$

(b) With $np = 100(0.70) = 70$ and $np(1-p) = 100(0.70)(0.30) = 21$, $P(75 \le X \le 80) = P(74.5 \le X \le 80.5) = P\left(\frac{74.5-70}{\sqrt{21}} \le \frac{X-70}{\sqrt{21}} \le \frac{80.5-70}{\sqrt{21}}\right) = P(0.98 \le Z \le 2.29) = 0.1525.$

4.3.11 Let $p = P$(person dies by chance in the three months following birthmonth) = $\frac{1}{4}$. Given that n = 747, np = 186.75, and $np(1 - p)$ = 140.06, $P(X \geq 344) = P(X \geq 343.5) =$
$P\left(\frac{X - 186.75}{\sqrt{140.06}} \geq \frac{343.5 - 186.75}{\sqrt{140.06}}\right) = P(Z \geq 13.25) = 0.0000$. The fact that the latter probability is so small strongly discredits the hypothesis that people die randomly with respect to their birthdays.

4.3.12 Let X = number of correct guesses (out of n = 1500 attempts). Since five choices were available for each guess (recall Figure 4.3.4), $p = P$(correct answer) = 1/5, if ESP is not a factor. Then $P(X \geq 326) = P(X \geq 325.5) = P\left(\frac{X - 1500(1/5)}{\sqrt{1500(1/5)(4/5)}} \geq \frac{325.5 - 300}{\sqrt{240}}\right) \doteq P(Z \geq 1.65) =$ 0.0495. Based on these results, there is certainly <u>some</u> evidence that ESP may be increasing the probability of a correct guess, but the magnitude of $P(X \geq 326)$ is not so small that it precludes the possibility that chance is the only operative factor.

4.3.13 No, the normal approximation is inappropriate because the values of n (= 10) and p (= 0.7) fail to satisfy the condition $n > 9\frac{p}{1-p} = 9\frac{0.7}{0.3} = 21$.

4.3.14 Let X = number of fans buying hot dogs. To be determined is the smallest value of c for which $P(X > c) \leq 0.20$. Assume that no one eats more than one hot dog. Then X is a binomial random variable with n = 42,200 and $p = P$(fan buys hot dog) = 0.38. Since np = 16,036 and $\sqrt{np(1-p)}$ = 99.7, $P(X > c) = 0.20 = P(X \geq c + 1) \doteq P\left(Z \geq \frac{c + 1 - \frac{1}{2} - 16{,}036}{99.7}\right)$.

But $P(Z \geq 0.8416) = 0.20$, so $0.8416 = \left(\frac{c + 1 - \frac{1}{2} - 16{,}036}{99.7}\right)$, from which it follows that

c = 16,119.

4.3.15 $P(|X - E(X)| \leq 5) = P(-5 \leq X - 100 \leq 5) = P\left(\frac{-5.5}{\sqrt{50}} \leq \frac{X - 100}{\sqrt{50}} \leq \frac{5.5}{\sqrt{50}}\right) \doteq$
$P(-0.78 \leq Z \leq 0.78) = 0.5646$.

For binomial data, the central limit theorem and DeMoivre-Laplace approximations differ only if the continuity correction is used in the DeMoivre-Laplace approximation.

4.3.16 Let X_i = face showing on ith die, $i = 1, 2, \ldots, 100$, and let $X = X_1 + X_2 + \ldots + X_{100}$. Following the approach taken in Example 3.9.5 gives $E(X) = 350$. Also, $\text{Var}(X_i) = E\left(X_i^2\right) - [E(X_i)]^2 =$

$\frac{1}{6}(1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2) - \left(3\frac{1}{2}\right)^2 = \frac{35}{12}$, so $\text{Var}(X) = \frac{3500}{12}$. By the central limit theorem,

then, $P(X > 370) = P(X \geq 371) = P(X \geq 370.5) = P\left(\frac{X - 350}{\sqrt{3500/12}} \geq \frac{370.5 - 350}{\sqrt{3500/12}}\right) \doteq P(Z \geq 1.20)$

$= 0.1151$.

4.3.17 For the given X, $E(X) = 5(18/38) + (-5)(20/38) = -10/38 = -0.263$.
$\text{Var}(X) = 25(18/38) + (25)(20/38) - (-10/38)^2 = 24.931$, $\sigma = 4.993$.

Then $P(X_1 + X_2 + \ldots + X_{100} > -50)$

$= P\left(\frac{X_1 + X_2 + \ldots + X_{100} - 100(-0.263)}{\sqrt{100}(4.993)} > \frac{-50 - 100(-0.263)}{10(4.993)}\right) \doteq P(Z > -0.47)$

$= 1 - 0.3192 = 0.6808$

4.3.18 If X_i is a Poisson random variable with parameter λ_i, then $E(X_i) = \text{Var}(X_i) = \lambda_i$. Let $X = X_1 + X_2 + \ldots + X_n$ be a sum of independent Poissons. Then $E(X) = \sum_{i=1}^{n} \lambda_i = \text{Var}(X)$. If $\lambda = \sum_{i=1}^{n} \lambda_i$, the ratio in Theorem 4.3.2 reduces to $\frac{X - \lambda}{\sqrt{\lambda}}$.

4.3.19 Let X = number of chips ordered next week. Given that $\lambda = E(X) = 50$, P(company is unable to fill orders) $= P(X \geq 61) = P(X \geq 60.5) = P\left(\frac{X - 50}{\sqrt{50}} \geq \frac{60.5 - 50}{\sqrt{50}}\right)$ $P(Z \geq 1.48) = 0.0694$.

4.3.20 Let X = number of leukemia cases diagnosed among 3000 observers. If $\lambda = E(X) = 3$,

$P(X \geq 8) = 1 - P(X \leq 7) = 1 - \sum_{k=0}^{7} \frac{e^{-3}3^k}{k!} = 1 - 0.9881 = 0.0119$. Using the central limit theorem,

$1 - P(X \leq 7) = 1 - P(X \leq 7.5) = 1 - P\left(\frac{X - 3}{\sqrt{3}} \leq \frac{7.5 - 3}{\sqrt{3}}\right) \doteq 1 - P(Z \leq 2.60) = 0.0047$. The approximation is not particularly good because λ is small. In general, if λ is less than 5, the normal approximation should not be used. Both analyses, though, suggest that the observer's risk of contracting leukemia was increased because of their exposure to the test.

4.3.21 No, only 84% of drivers are likely to get at least 25,000 miles on the tires. If X denotes the mileage obtained on a set of Econo-Tires, $P(X \geq 25{,}000) =$

$P\left(\frac{X - 30{,}000}{5000} \geq \frac{25{,}000 - 30{,}000}{5000}\right) = P(Z \geq -1.00) = 0.8413$.

4.3.22 Let Y denote a child's IQ. Then P(child needs special services) $= P(Y < 80) + P(Y > 135) =$
$P\left(\frac{Y-100}{16} < \frac{80-100}{16}\right) + P\left(\frac{Y-100}{16} > \frac{135-100}{16}\right) = P(Z < -1.25) + P(Z > 2.19) =$
$0.1056 + 0.0143 = 0.1199$. It follows that $1400 \times 0.1199 \times \$1750 = \$293{,}755$ should be added to Westbank's special ed budget.

4.3.23 Let Y = donations collected tomorrow. Given that $\mu = \$20{,}000$ and $\sigma = \$5{,}000$,
$P(Y > \$30{,}000) = P\left(\frac{Y - \$20{,}000}{\$5{,}000} > \frac{\$30{,}000 - \$20{,}000}{\$5{,}000}\right) = P(Z > 2.00) = 0.0228.$

4.3.24 Let Y = pregnancy duration (in days). Ten months and five days is equivalent to 310 days. The credibility of San Diego Reader's claim hinges on the magnitude of $P(Y \geq 310)$—the smaller that probability is, the less believable her explanation becomes. Given that $\mu = 266$ and $\sigma = 16$, $P(Y \geq 310) = P\left(\frac{Y-266}{16} \geq \frac{310-266}{16}\right) = P(Z \geq 2.75) = 0.0030$. While the latter does not rule out the possibility that San Diego Reader is telling the truth, pregnancies lasting 310 or more days are extremely unlikely.

4.3.25 **(a)** Let Y_1 and Y_2 denote the scores made by a random nondelinquent and delinquent, respectively. Then $E(Y_1) = 60$ and $\text{Var}(Y_1) = 10^2$; also, $E(Y_2) = 80$ and $\text{Var}(Y_2) = 5^2$. Since 75 is the cutoff between teenagers classified as delinquents or nondelinquents,
P(nondelinquent is misclassified as delinquent) $= P(Y_1 > 75) = P\left(Z > \frac{75-60}{10}\right) = 0.0668.$
Similarly, P(delinquent is misclassified as nondelinquent) $= P(Y_2 < 75) = P\left(Z < \frac{75-80}{5}\right) =$
$0.1587.$

4.3.26 Let Y denote the cross-sectional area of a tube. Then $p = P$(tube does not fit properly) $=$
$P(Y < 12.0) + P(Y > 13.0) = 1 - P(12.0 \leq Y \leq 13.0) = 1 - P\left(\frac{12.0-12.5}{0.2} \leq \frac{Y-12.5}{0.2} \leq \frac{13.0-12.5}{0.2}\right) =$
$1 - P(-2.50 \leq Z \leq 2.50) = 1 - 0.9876 = 0.0124$. Let X denote the number of tubes (out of 1000) that will not fit. Since X is a binomial random variable, $E(X) = np = 1000(0.0124) = 12.4$.

4.3.27 Let Y = freshman's verbal SAT score. Given that $\mu = 565$ and $\sigma = 75$, $P(Y > 660) =$
$P\left(\frac{Y-565}{75} > \frac{660-565}{75}\right) = P(Z > 1.27) = 0.1020$. It follows that the expected <u>number</u> doing better is 4250(0.1020), or 434.

4.3.28 Let A^* and B^* denote the lowest A and the lowest B, respectively. Since the top 20% of the grades will be A's, $P(Y < A^*) = 0.80$, where Y denotes a random student's score. Equivalently,
$P\left(Z < \frac{A^*-70}{12}\right) = 0.80$. From Appendix Table A.1, though, $P(Z < 0.84) = 0.7995 \doteq 0.80$.
Therefore, $0.84 = \frac{A^*-70}{12}$, which implies that $A^* = 80$.

Similarly, $P(Y < B^*) = 0.54 = P\left(Z < \frac{B^* - 70}{12}\right)$. But $P(Z < 0.10) = 0.5398 \doteq 0.54$,

so $0.10 = \frac{B^* - 70}{12}$, implying that $B^* = 71$.

4.3.29 If $P(20 \le Y \le 60) = 0.50$, then $P\left(\frac{20-40}{\sigma} \le \frac{Y-40}{\sigma} \le \frac{60-40}{\sigma}\right) = 0.50 = P\left(\frac{-20}{\sigma} \le Z \le \frac{20}{\sigma}\right)$. But $P(-0.67 \le Z \le 0.67) = 0.4972 \doteq 0.50$, which implies that $0.67 = \frac{20}{\sigma}$. The desired value for σ, then, is $\frac{20}{0.67}$, or 29.85.

4.3.30 Let Y = a random 18-year-old woman's weight. Since $\mu = \frac{103.5 + 144.5}{2} = 124$,

$P(103.5 \le Y \le 144.5) = 0.80 = P\left(\frac{103.5-124}{\sigma} \le \frac{Y-124}{\sigma} \le \frac{144.5-124}{\sigma}\right) = P\left(\frac{-20.5}{\sigma} \le Z \le \frac{20.5}{\sigma}\right)$.

According to Appendix Table A.1, $P(-1.28 \le Z \le 1.28) \doteq 0.80$, so $\frac{20.5}{\sigma} = 1.28$, implying that $\sigma = 16.0$ lbs.

4.3.31 Let Y = analyzer reading for driver whose true blood alcohol concentration is 0.9. Then $P(\text{analyzer mistakenly shows driver to be sober}) = P(Y < 0.08) = P\left(\frac{Y-0.9}{0.004} < \frac{0.08-0.09}{0.004}\right) = P(Z < -2.50) = 0.0062$. The "0.075%" driver should ask to take the test twice. The "0.09%" driver has a greater chance of not being charged by taking the test only once. As, n the number of times the test taken, increases, the precision of the average reading increases. It is to the sober driver's advantage to have a reading as precise as possible; the opposite is true for the drunk driver.

4.3.32 The normed score for Michael is $\frac{75-62.0}{7.6} = 1.71$; the normed score for Laura is $\frac{92-76.3}{10.8} = 1.45$. So, even though Laura made 17 points higher on the test, the company would be committed to hiring Michael.

4.3.33 By the first corollary to Theorem 4.3.3, $P(\bar{Y} > 103) = P\left(\frac{\bar{Y}-100}{16/\sqrt{9}} > \frac{103-100}{16/\sqrt{9}}\right) = P(Z > 0.56) = 0.2877$. For any arbitrary Y_i, $P(Y_i > 103) = P\left(\frac{Y_i - 100}{16} > \frac{103-100}{16}\right) = P(Z > 0.19) = 0.4247$.

Let X = number of Y_i's that exceed 103. Since X is a binomial random variable with $n = 9$ and $p = P(Y_i > 103) = 0.4247$, $P(X = 3) = \binom{9}{3}(0.4247)^3(0.5753)^6 = 0.23$.

4.3.34 If $P(1.9 \le \bar{Y} \le 2.1) \ge 0.99$, then $P\left(\frac{1.9-2}{2/\sqrt{n}} \le Z \le \frac{2.1-2}{2/\sqrt{n}}\right) \ge 0.99$. But $P(-2.58 \le Z \le 2.58)$ $\doteq 0.99$, so $2.58 = \frac{2.1-2}{2/\sqrt{n}}$, which implies that $n = 2663$.

4.3.35 Let Y_i = resistance of ith resistor, $i = 1, 2, 3$, and let $Y = Y_1 + Y_2 + Y_3$ = circuit resistance. By the first corollary to Theorem 4.3.3, $E(Y) = 6 + 6 + 6 = 18$ and $\text{Var}(Y) = (0.3)^2 + (0.3)^2 + (0.3)^2 = 0.27$. Therefore, $P(Y > 19) = P\left(\frac{Y-18}{\sqrt{0.27}} > \frac{19-18}{\sqrt{0.27}}\right) = P(Z > 1.92) = 0.0274$. Suppose $P(Y > 19)$ $= 0.005 = P\left(Z > \frac{19-18}{\sqrt{3\sigma^2}}\right)$. From Appendix Table A.1, $P(Z > 2.58) \doteq 0.005$, so $2.58 = \frac{19-18}{\sqrt{3\sigma^2}}$, which implies that the minimum "precision" of the manufacturing process would have to be $\sigma =$ 0.22 ohms.

4.3.36 Let Y_P and Y_C denote a random piston diameter and cylinder diameter, respectively. Then P(pair needs to be reworked) = $P(Y_P > Y_C) = P(Y_P - Y_C > 0)$

$$= P\left(\frac{Y_P - Y_C - (40.5-41.5)}{\sqrt{(0.3)^2+(0.4)^2}} > \frac{0-(40.5-41.5)}{\sqrt{(0.3)^2+(0.4)^2}}\right) = P(Z > 2.00) = 0.0228, \text{ or } 2.28\%.$$

4.3.37 $M_{\bar{Y}}(t) = M_{Y_1+\ldots Y_n}\left(\frac{t}{n}\right) = \prod_{i=1}^{n} M_{Y_i}\left(\frac{t}{n}\right) = \prod_{i=1}^{n} e^{\mu t/n + \sigma^2 t^2/2n^2} = e^{\mu t + \sigma^2 t^2/2n}$, but the latter is the moment-generating function for a normal random variable whose mean is μ and whose variance is σ^2/n. Similarly, if $Y = a_1Y_1 + \ldots + a_nY_n$, $M_Y(t) = \prod_{i=1}^{n} M_{a_iY_i}(t) = \prod_{i=1}^{n} M_{Y_i}(a_it) =$ $\prod_{i=1}^{n} e^{\mu_i a_i t + \sigma_i^2 a_i^2 t^2/2} = e^{\sum_{i=1}^{n} a_i\mu_i t + \sum_{i=1}^{n} a_i^2\sigma_i^2 t^2/2}$. By inspection, Y has the moment-generating function of a normal random variable for which $E(Y) = \sum_{i=1}^{n} a_i\mu_i$ and $\text{Var}(Y) = \sum_{i=1}^{n} a_i^2\sigma_i^2$.

4.3.38 $P(\bar{Y} \ge \bar{Y}^*) = P(\bar{Y} - \bar{Y}^* \ge 0)$, where $E(\bar{Y} - \bar{Y}^*) = E(\bar{Y}) - E(\bar{Y}^*) = 2 - 1 = 1$.

Also, $\text{Var}(\bar{Y} - \bar{Y}^*) = \text{Var}(\bar{Y}) - \text{Var}(\bar{Y}^*) = \frac{2^2}{9} + \frac{1^2}{4} = \frac{25}{36}$, because $\bar{Y}$ and $\bar{Y}^*$ are independent.

Therefore, $P(\bar{Y} \ge \bar{Y}^*) = P\left(\frac{\bar{Y} - \bar{Y}^* - 1}{\sqrt{25/36}} \ge \frac{0-1}{\sqrt{25/36}}\right) = P(Z \ge -1.20) = 0.8849$

Section 4.4: The Geometric Distribution

4.4.1 Let $p = P$(return is audited in a given year) = 0.30 and let X = year of first audit. Then P(Jody escapes detection for at least 3 years) = $P(X \geq 4) = 1 - P(X \leq 3)$ =

$$1 - \sum_{k=1}^{3}(0.70)^{k-1}(0.30) = 0.343.$$

4.4.2 If X = attempt at which license is awarded and $p = P$(driver passes test on any given attempt) = 0.10, then $p_X(k) = (0.90)^{k-1}(0.10)$, $k = 1, 2, \ldots$; $E(X) = \frac{1}{p} = \frac{1}{0.10} = 10$.

4.4.3 No, the expected frequencies (= $5 \cdot p_X(k)$) differ considerably from the observed frequencies, especially for small values of k. The observed number of 1's, for example, is 4, while the expected number is 12.5.

k	Obs. Freq.	$p_X(k) = \left(\frac{3}{4}\right)^{k-1}\left(\frac{1}{4}\right)$	$50 \cdot p_X(k)$ = Exp. freq.
1	4	0.2500	12.5
2	13	0.1875	9.4
3	10	0.1406	7.0
4	7	0.1055	5.3
5	5	0.0791	4.0
6	4	0.0593	3.0
7	3	0.0445	2.2
8	3	0.0334	1.7
9+	1	0.1001	5.0
	50	1.0000	50.0

4.4.4 If $p = P$(child is a girl) and X = birth order of first girl, then $E(X) = \frac{1}{p} = \frac{1}{\frac{1}{2}} = 2$. Barring any medical restrictions, it would not be unreasonable to model the appearance of a couple's first girl (or boy) by the geometric probability function. The most appropriate value for p, though, would not be exactly $\frac{1}{2}$ (although census figures indicate that it would be close to $\frac{1}{2}$).

4.4.5 $F_X(t) = P(X \leq t) = p\sum_{s=0}^{[t]}(1-p)^s$. But $\sum_{s=0}^{[t]}(1-p)^s = \frac{1-(1-p)^{[t]}}{1-(1-p)} = \frac{1-(1-p)^{[t]}}{p}$, and the result follows.

4.4.6 Let X = roll on which sum of 4 appears for first time. Since $p = P(\text{sum} = 4) = \frac{3}{216}$. $p_X(k) = \left(\frac{213}{216}\right)^{k-1} \cdot \frac{3}{216}$, $k = 1, 2, \ldots$ Using the expression for $F_X(k)$ given in Question 4.4.5, we can write $P(65 \le X \le 75) = F_X(75) - F_X(64) = 1 - \left(1-\frac{3}{216}\right)^{[75]} - \left(1-\left(1-\frac{3}{216}\right)^{[64]}\right) = \left(\frac{213}{216}\right)^{64} - \left(\frac{213}{216}\right)^{75} = 0.058$.

4.4.7 $P(n \le Y \le n + 1) = \int_n^{n+1} \lambda e^{-\lambda y}\,dy = (1-e^{-\lambda y})\Big|_n^{n+1} = e^{-\lambda n} - e^{-\lambda(n+1)} = e^{-\lambda n}(1-e^{-\lambda})$
Setting $p = 1 - e^{-\lambda}$ gives $P(n \le Y \le n + 1) = (1 - p)^n p$.

4.4.8 Let the random variable X^* denote the number of trials preceding the first success. By inspection, $p_{X^*}(t) = (1 - p)^k p$, $k = 0, 1, 2, \ldots$ Also, $M_{X^*}(t) = \sum_{k=0}^{\infty} e^{tk} \cdot (1-p)^k p = p\sum_{k=0}^{\infty}[(1-p)e^t]^k = p \cdot \left(\frac{1}{1-(1-p)e^t}\right) = \frac{p}{1-(1-p)e^t}$. Let X denote the geometric random variable defined in Theorem 4.4.1. Then $X^* = X - 1$, and $M_{X^*}(t) = e^{-t}M_X(t) = e^{-t} \cdot \frac{pe^t}{1-(1-p)e^t} = \frac{p}{1-(1-p)e^t}$.

4.4.9 $M_X(t) = pe^t[1 - (1 - p)e^t]^{-1}$, so $M_X^{(1)}(t) = pe^t(-1)[1 - (1 - p)e^t]^{-2} \cdot (-(1 - p)e^t) +$ $[1 - (1 - p)e^t]^{-1}pe^t$. Setting $t = 0$ gives $M_X^{(1)}(0) = E(X) = \frac{1}{p}$. Similarly, $M_X^{(2)}(t) =$ $p(1 - p)e^{2t}(-2)[1 - (1 - p)e^t]^{-3} \cdot (-(1 - p)e^t) + [1 - (1 - p)e^t]^{-2}p(1 - p)e^{2t} \cdot 2 +$ $[1 - (1 - p)e^t]^{-1}pe^t + pe^t(-1)[1 - (1 - p)e^t]^{-2} \cdot (-(1 - p)e^t)$ and $M_X^{(2)}(0) = E(X^2) = \frac{2-p}{p^2}$.
Therefore, $\text{Var}(X) = E(X^2) - [E(X)]^2 = \frac{2-p}{p^2} - \left(\frac{1}{p}\right)^2 = \frac{1-p}{p^2}$

4.4.10 No, because $M_X(t) = M_{X_1}(t) \cdot M_{X_2}(t)$ does not have the form of a geometric moment-generating function.

4.4.11 Let $M_X^*(t) = E(t^X) = \sum_{k=1}^{\infty} t^k \cdot (1-p)^{k-1} p = \frac{p}{1-p}\sum_{k=1}^{\infty}[t(1-p)]^k = \frac{p}{1-p}\sum_{k=0}^{\infty}[t(1-p)]^k - \frac{p}{1-p} =$ $\frac{p}{1-p}\left[\frac{1}{1-t(1-p)}\right] - \frac{p}{1-p} = \frac{pt}{1-t(1-p)}$ = factorial moment-generating function for X.
Then $M_X^{*(1)}(t) = pt(-1)[1 - t(1 - p)]^{-2}(-(1 - p)) + [1 - t(1 - p)]^{-1}p = \frac{p}{[1-t(1-p)]^2}$.

When $t = 1$, $M_X^{*(1)}(1) = E(X) = \dfrac{1}{p}$. Also, $M_X^{*(2)}(t) = \dfrac{2p(1-p)}{[1-t(1-p)]^3}$ and $M_X^{*(2)}(1) = \dfrac{2-2p}{p^2} =$ $E[X(X-1)] = E(X^2) - E(X)$. Therefore, $\text{Var}(X) = E(X^2) - [E(X)]^2 = \dfrac{2-2p}{p^2} + \dfrac{1}{p} - \left(\dfrac{1}{p}\right)^2$ $= \dfrac{1-p}{p^2}$.

Section 4.5: The Negative Binomial Distribution

4.5.1 Let X = number of houses needed to achieve fifth invitation. If $p = P$(saleswoman receives invitation at a given house) = 0.30, $p_X(k) = \binom{k-1}{4}(0.30)^4(0.70)^{k-1-4}(0.30)$, $k = 5, 6, \ldots$ and $P(X < 8) = P(5 \le X \le 7) = \sum_{k=5}^{7}\binom{k-1}{4}(0.30)^5(0.70)^{k-5} = 0.029$.

4.5.2 Let $p = P$(missile scores direct hit) = 0.30. Then P(target will be destroyed by seventh missile fired) = P(exactly three direct hits occur among first six missiles <u>and</u> seventh missile scores direct hit) = $\binom{6}{3}(0.30)^3(0.70)^3(0.30) = 0.056$.

4.5.3 Darryl might have actually done his homework, but there is reason to suspect that he did not. Let the random variable X denote the toss where a head appears for the second time. Then $p_X(k) = \binom{k-1}{1}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^{k-2}$, $k = 2, 3, \ldots$, but that particular model fits the data almost perfectly, as the table shows. Agreement this good is often an indication that the data have been fabricated.

k	$p_X(k)$	Obs. freq.	Exp. freq.
2	1/4	24	25
3	2/8	26	25
4	3/16	19	19
5	4/32	13	12
6	5/64	8	8
7	6/128	5	5
8	7/256	3	3
9	8/512	1	2
10	9/1024	1	1

4.5.4 Let $p = P$(defective is produced by improperly adjusted machine) = 0.15. Let X = item at which machine is readjusted. Then $p_X(k) = \binom{k-1}{2}(0.15)^2(0.85)^{k-1-2}(0.15) = \binom{k-1}{2}(0.15)^3(0.85)^{k-3}$, $k = 3, 4, \ldots$ It follows that $P(X \ge 5) = 1 - P(X \le 4)$ $= 1 - [P(X = 3) + P(X = 4)] = 0.988$ and $E(X) = \dfrac{3}{0.15} = 20$.

4.5.5 $E(X) = \sum_{k=r}^{\infty} k\binom{k-1}{r-1}p^r(1-p)^{k-r} = \frac{r}{p}\sum_{k=r}^{\infty}\binom{k}{r}p^{r+1}(1-p)^{k-r} = \frac{r}{p}$.

4.5.6 Let Y denote the number of trials to get the rth success, and let X denote the number of trials in excess of r to get the rth success. Then $X = Y - r$. Substituting into Theorem 4.5.1 gives
$p_X(k) = \binom{k+r-1}{r-1}p^r(1-p)^k = \binom{k+r-1}{k}p^r(1-p)^k$, $k = 0, 1, 2, \ldots$

4.5.7 Here $X = Y - r$, where Y has the negative binomial distribution as described in Theorem 4.5.1. Using the properties (1), (2), and (3) given by the theorem, we can write $E(X) = E(Y - r)$
$= E(Y) - E(r) = \frac{r}{p} - r = \frac{r(1-p)}{p}$ and $\text{Var}(X) = \text{Var}(Y - r) = \text{Var}(Y) + \text{Var}(r)$
$= \frac{r(1-p)}{p^2} + 0 = \frac{r(1-p)}{p^2}$. Also, $M_X(t) = M_{Y-r}(t) = e^{-rt}M_Y(t) = e^{-rt}\left[\frac{pe^t}{1-(1-p)e^t}\right]^r = \left[\frac{p}{1-(1-p)e^t}\right]^r$.

4.5.8 For each X_i, $M_{X_i}(t) = \left[\frac{(4/5)e^t}{1-(1-4/5)e^t}\right]^3$, $i = 1, 2, 3$. If $X = X_1 + X_2 + X_3$, it follows that

$M_X(t) = \prod_{i=1}^{3} M_{X_i}(t) = \left[\frac{(4/5)e^t}{1-(1-4/5)e^t}\right]^9$, which implies that $p_X(k) = \binom{k-1}{8}\left(\frac{4}{5}\right)^9\left(\frac{1}{5}\right)^{k-9}$,

$k = 9, 10, \ldots$ Then $P(10 \le X \le 12) = \sum_{k=10}^{12} p_X(k) = 0.66$.

4.5.9 $M_X^{(1)}(t) = r\left[\frac{pe^t}{1-(1-p)e^t}\right]^{r-1}[pe^t[1-(1-p)e^t]^{-2}(1-p)e^t + [1-(1-p)e^t]^{-1}pe^t]$. When $t = 0$,

$M_X^{(1)}(0) = E(X) = r\left[\frac{p(1-p)}{p^2} + \frac{p}{p}\right] = \frac{r}{p}$.

4.5.10 $M_X(t) = \prod_{i=1}^{k} M_{X_i}(t) = \prod_{i=1}^{k}\left[\frac{pe^t}{1-(1-p)e^t}\right]^{r_i} = \left[\frac{pe^t}{1-(1-p)e^t}\right]^{r^*}$, where $r^* = \sum_{i=1}^{k} r_i$. Also,

$E(X) = E(X_1 + X_2 + \ldots + X_k) = E(X_1) + E(X_2) + \ldots + E(X_k) = \sum_{i=1}^{k}\frac{r_i}{p} = \frac{r^*}{p}$ and

$\text{Var}(X) = \sum_{i=1}^{k}\text{Var}(X_i) = \sum_{i=1}^{k}\frac{r_i(1-p)}{p^2} = \frac{r^*(1-p)}{p^2}$.

Section 4.6: The Gamma Distribution

4.6.1 Let Y_i = lifetime of ith gauge, i = 1, 2, 3. By assumption, $f_{Y_i}(y) = 0.001e^{-0.001y}$, $y > 0$. Define the random variable $Y = Y_1 + Y_2 + Y_3$ to be the lifetime of the system. By Theorem 4.6.1, $f_Y(y) = \dfrac{(0.001)^3}{2} y^2 e^{-0.001y}$, $y > 0$.

4.6.2 The time until the 24th breakdown is a gamma random variable with parameters $r = 24$ and $\lambda = 3$. The mean of this random variable is $r/\lambda = 24/3 = 8$ months.

4.6.3 If $E(Y) = \dfrac{r}{\lambda} = 1.5$ and $\text{Var}(Y) = \dfrac{r}{\lambda^2} = 0.75$, then $r = 3$ and $\lambda = 2$, which makes $f_Y(y) = 4y^2e^{-2y}$, $y > 0$. Then $P(1.0 \le Y_i \le 2.5) = \int_{1.0}^{2.5} 4y^2 e^{-2y}dy = 0.55$. Let X = number of Y_i's in the interval (1.0, 2.5). Since X is a binomial random variable with $n = 100$ and $p = 0.55$, $E(X) = np = 55$.

4.6.4 $f_{\lambda Y}(y) = \dfrac{1}{\lambda} f_Y(y/\lambda) = \dfrac{1}{\lambda}\dfrac{\lambda^r}{\Gamma(r)}\left(\dfrac{y}{\lambda}\right)^{r-1} e^{-\lambda(y/\lambda)} = \dfrac{1}{\Gamma(r)} y^{r-1}e^{-y}$

4.6.5 To find the maximum of the function $f_Y(y) = \dfrac{\lambda^r}{\Gamma(r)} y^{r-1}e^{-\lambda y}$, differentiate it with respect to y and set it equal to 0; that is

$$\frac{df_Y(y)}{dy} = \frac{d}{dy}\frac{\lambda^r}{\Gamma(r)} y^{r-1}e^{-\lambda y} = \frac{\lambda^r}{\Gamma(r)}[(r-1)y^{r-2}e^{-\lambda y} - \lambda y^{r-1}e^{-\lambda y}] = 0$$

This implies $\dfrac{\lambda^r}{\Gamma(r)} y^{r-2}e^{-\lambda y}[(r-1) - \lambda y] = 0$, whose solution is $y_{\text{mode}} = \dfrac{r-1}{\lambda}$. Since the derivative is positive for $y < y_{\text{mode}}$, and negative for $y > y_{\text{mode}}$, then there is a maximum.

4.6.6 Let Z be a standard normal random variable. Then $E(Z^2) = \dfrac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} z^2 e^{-z^2/2}dz$ $= \sqrt{\dfrac{2}{\pi}}\int_0^{\infty} z^2 e^{-z^2/2}dz = 1$. Let $y = z^2$. Then $E(Z^2) = \dfrac{2}{\sqrt{\pi}}\Gamma\left(1+\dfrac{1}{2}\right) = \dfrac{2}{\sqrt{\pi}}\left(\dfrac{1}{2}\right)\Gamma\left(\dfrac{1}{2}\right)$, which implies that $\Gamma\left(\dfrac{1}{2}\right) = \sqrt{\pi}$.

4.6.7 $\Gamma\left(\dfrac{7}{2}\right) = \dfrac{5}{2}\Gamma\left(\dfrac{5}{2}\right) = \dfrac{5}{2}\dfrac{3}{2}\Gamma\left(\dfrac{3}{2}\right) = \dfrac{5}{2}\dfrac{3}{2}\dfrac{1}{2}\Gamma\left(\dfrac{1}{2}\right) = \dfrac{15}{8}\Gamma\left(\dfrac{1}{2}\right)$ by Theorem 4.6.2, part 2. Further, $\Gamma\left(\dfrac{1}{2}\right) = \sqrt{\pi}$ by Question 4.6.6.

4.6.8 $E(Y^m) = \int_0^\infty y^m \cdot \frac{\lambda^r}{(r-1)!} y^{r-1} e^{-\lambda y} dy = \int_0^\infty \frac{\lambda^r}{(r-1)!} y^{m+r-1} e^{-\lambda y} dy$

$= \frac{(m+r-1)!}{\lambda^m (r-1)!} \int_0^\infty \frac{\lambda^{m+r}}{(m+r-1)!} y^{m+r-1} e^{-\lambda y} dy = \frac{(m+r-1)!}{\lambda^m (r-1)!}$.

4.6.9 Write the gamma moment-generating function in the form $M_Y(t) = (1 - t/\lambda)^{-r}$. Then $M_Y^{(1)}(t) = -r(1 - t/\lambda)^{-r-1}(-1/\lambda) = (r/\lambda)(1 - t/\lambda)^{-r-1}$ and $M_Y^{(2)}(t) = (r/\lambda)(-r-1)(1 - t/\lambda)^{r-2} \cdot (-1/\lambda)$ $= (r/\lambda^2)(r+1)(1 - t/\lambda)^{-r-2}$. Therefore, $E(Y) = M_Y^{(1)}(0) = \frac{r}{\lambda}$ and $\text{Var}(Y) = M_Y^{(2)}(0) - \left[M_Y^{(1)}(0)\right]^2$ $= \frac{r(r+1)}{\lambda^2} - \frac{r^2}{\lambda^2} = \frac{r}{\lambda^2}$.

4.6.10 $M_Y(t) = (1 - t/\lambda)^{-r}$ so $M_Y^{(1)}(t) = \frac{d}{dt}(1-t/\lambda)^{-r} = r(1-t/\lambda)^{-r-1}(-1/\lambda) = \frac{r}{\lambda}(1-t/\lambda)^{-r-1}$ and

$M_Y^{(2)}(t) = \frac{d}{dt}\frac{r}{\lambda}(1-t/\lambda)^{-r-1} = \frac{r}{\lambda}(-r-1)(1-t/\lambda)^{-r-2}(-1/\lambda) = \frac{r(r+1)}{\lambda^2}(1-t/\lambda)^{-r-2}$

For an arbitrary integer $m \geq 2$, we can generalize the above to see that

$M_Y^{(m)}(t) = \frac{r(r+1)...(r+m-1)}{\lambda^m}(1-t/\lambda)^{-r-m}$. Then $E(Y^m) = M_Y^{(m)}(0) = \frac{r(r+1)...(r+m-1)}{\lambda^m}$.

But note that $\frac{r(r+1)...(r+m-1)}{\lambda^m} = \frac{\Gamma(r+m)}{\Gamma(r)\lambda^m}$. The right hand side of the equation is equal to the expression in Question 4.6.8 when r is an integer.