Chapter 3

Common Families of Distributions

3.1 The pmf of X is $f(x) = \frac{1}{N_1-N_0+1}$, $x = N_0, N_0+1, \ldots, N_1$. Then

$$\begin{aligned} \mathrm{E}X &= \sum_{x=N_0}^{N_1} x \frac{1}{N_1-N_0+1} = \frac{1}{N_1-N_0+1}\left(\sum_{x=1}^{N_1} x - \sum_{x=1}^{N_0-1} x\right) \\ &= \frac{1}{N_1-N_0+1}\left(\frac{N_1(N_1+1)}{2} - \frac{(N_0-1)(N_0-1+1)}{2}\right) \\ &= \frac{N_1+N_0}{2}. \end{aligned}$$

Similarly, using the formula for $\sum_1^N x^2$, we obtain

$$\begin{aligned} \mathrm{E}x^2 &= \frac{1}{N_1-N_0+1}\left(\frac{N_1(N_1+1)(2N_1+1) - N_0(N_0-1)(2N_0-1)}{6}\right) \\ \mathrm{Var}X &= \mathrm{E}X^2 - \mathrm{E}X = \frac{(N_1-N_0)(N_1-N_0+2)}{12}. \end{aligned}$$

3.2 Let X = number of defective parts in the sample. Then $X \sim \text{hypergeometric}(N=100, M, K)$ where M = number of defectives in the lot and K = sample size.

a. If there are 6 or more defectives in the lot, then the probability that the lot is accepted $(X = 0)$ is at most

$$P(X=0 \mid M=100, N=6, K) = \frac{\binom{6}{0}\binom{94}{K}}{\binom{100}{K}} = \frac{(100-K)\cdot \cdots \cdot (100-K-5)}{100\cdot \cdots \cdot 95}.$$

By trial and error we find $P(X=0) = .10056$ for $K=31$ and $P(X=0) = .09182$ for $K=32$. So the sample size must be at least 32.

b. Now $P(\text{accept lot}) = P(X = 0 \text{ or } 1)$, and, for 6 or more defectives, the probability is at most

$$P(X=0 \text{ or } 1 \mid M=100, N=6, K) = \frac{\binom{6}{0}\binom{94}{K}}{\binom{100}{K}} + \frac{\binom{6}{1}\binom{94}{K-1}}{\binom{100}{K}}.$$

By trial and error we find $P(X=0 \text{ or } 1) = .10220$ for $K=50$ and $P(X=0 \text{ or } 1) = .09331$ for $K=51$. So the sample size must be at least 51.

3.3 In the seven seconds for the event, no car must pass in the last three seconds, an event with probability $(1-p)^3$. The only occurrence in the first four seconds, for which the pedestrian does not wait the entire four seconds, is to have a car pass in the first second and no other car pass. This has probability $p(1-p)^3$. Thus the probability of waiting exactly four seconds before starting to cross is $[1-p(1-p)^3](1-p)^3$.

3.5 Let X = number of effective cases. If the new and old drugs are equally effective, then the probability that the new drug is effective on a case is .8. If the cases are independent then $X \sim$ binomial(100, .8), and

$$P(X \geq 85) = \sum_{x=85}^{100} \binom{100}{x} .8^x .2^{100-x} = .1285.$$

So, even if the new drug is no better than the old, the chance of 85 or more effective cases is not too small. Hence, we cannot conclude the new drug is better. Note that using a normal approximation to calculate this binomial probability yields $P(X \geq 85) \approx P(Z \geq 1.125) = .1303$.

3.7 Let $X \sim$ Poisson(λ). We want $P(X \geq 2) \geq .99$, that is,

$$P(X \leq 1) = e^{-\lambda} + \lambda e^{-\lambda} \leq .01.$$

Solving $e^{-\lambda} + \lambda e^{-\lambda} = .01$ by trial and error (numerical bisection method) yields $\lambda = 6.6384$.

3.8 a. We want $P(X > N) < .01$ where $X \sim$ binomial(1000, 1/2). Since the 1000 customers choose randomly, we take $p = 1/2$. We thus require

$$P(X > N) = \sum_{x=N+1}^{1000} \binom{1000}{x} \left(\frac{1}{2}\right)^x \left(1 - \frac{1}{2}\right)^{1000-x} < .01$$

which implies that

$$\left(\frac{1}{2}\right)^{1000} \sum_{x=N+1}^{1000} \binom{1000}{x} < .01.$$

This last inequality can be used to solve for N, that is, N is the smallest integer that satisfies

$$\left(\frac{1}{2}\right)^{1000} \sum_{x=N+1}^{1000} \binom{1000}{x} < .01.$$

The solution is $N = 537$.

b. To use the normal approximation we take $X \sim$ n(500, 250), where we used $\mu = 1000(\frac{1}{2}) = 500$ and $\sigma^2 = 1000(\frac{1}{2})(\frac{1}{2}) = 250$.Then

$$P(X > N) = P\left(\frac{X - 500}{\sqrt{250}} > \frac{N - 500}{\sqrt{250}}\right) < .01$$

thus,

$$P\left(Z > \frac{N - 500}{\sqrt{250}}\right) < .01$$

where $Z \sim$ n(0, 1). From the normal table we get

$$P(Z > 2.33) \approx .0099 < .01 \quad \Rightarrow \quad \frac{N - 500}{\sqrt{250}} = 2.33$$
$$\Rightarrow \quad N \approx 537.$$

Therefore, each theater should have at least 537 seats, and the answer based on the approximation equals the exact answer.

3.9 a. We can think of each one of the 60 children entering kindergarten as 60 independent Bernoulli trials with probability of success (a twin birth) of approximately $\frac{1}{90}$. The probability of having 5 or more successes approximates the probability of having 5 or more sets of twins entering kindergarten. Then $X \sim \text{binomial}(60, \frac{1}{90})$ and

$$P(X \geq 5) = 1 - \sum_{x=0}^{4} \binom{60}{x} \left(\frac{1}{90}\right)^x \left(1 - \frac{1}{90}\right)^{60-x} = .0006,$$

which is small and may be rare enough to be newsworthy.

b. Let X be the number of elementary schools in New York state that have 5 or more sets of twins entering kindergarten. Then the probability of interest is $P(X \geq 1)$ where $X \sim \text{binomial}(310,.0006)$. Therefore $P(X \geq 1) = 1 - P(X = 0) = .1698$.

c. Let X be the number of States that have 5 or more sets of twins entering kindergarten during any of the last ten years. Then the probability of interest is $P(X \geq 1)$ where $X \sim \text{binomial}(500, .1698)$. Therefore $P(X \geq 1) = 1 - P(X = 0) = 1 - 3.90 \times 10^{-41} \approx 1$.

3.11 a.

$$\begin{aligned} &\lim_{M/N \to p, M \to \infty, N \to \infty} \frac{\binom{M}{x}\binom{N-M}{K-x}}{\binom{N}{K}} \\ &= \frac{K!}{x!(K-x)!} \lim_{M/N \to p, M \to \infty, N \to \infty} \frac{M!(N-M)!(N-K)!}{N!(M-x)!(N-M-(K-x))!} \end{aligned}$$

In the limit, each of the factorial terms can be replaced by the approximation from Stirling's formula because, for example,

$$M! = (M!/(\sqrt{2\pi}M^{M+1/2}e^{-M}))\sqrt{2\pi}M^{M+1/2}e^{-M}$$

and $M!/(\sqrt{2\pi}M^{M+1/2}e^{-M}) \to 1$. When this replacement is made, all the $\sqrt{2\pi}$ and exponential terms cancel. Thus,

$$\begin{aligned} &\lim_{M/N \to p, M \to \infty, N \to \infty} \frac{\binom{M}{x}\binom{N-M}{K-x}}{\binom{N}{K}} \\ &= \binom{K}{x} \lim_{M/N \to p, M \to \infty, N \to \infty} \frac{M^{M+1/2}(N-M)^{N-M+1/2}(N-K)^{N-K+1/2}}{N^{N+1/2}(M-x)^{M-x+1/2}(N-M-K+x)^{N-M-(K-x)+1/2}}. \end{aligned}$$

We can evaluate the limit by breaking the ratio into seven terms, each of which has a finite limit we can evaluate. In some limits we use the fact that $M \to \infty$, $N \to \infty$ and $M/N \to p$ imply $N - M \to \infty$. The first term (of the seven terms) is

$$\lim_{M \to \infty} \left(\frac{M}{M-x}\right)^M = \lim_{M \to \infty} \frac{1}{\left(\frac{M-x}{M}\right)^M} = \lim_{M \to \infty} \frac{1}{\left(1 + \frac{-x}{M}\right)^M} = \frac{1}{e^{-x}} = e^x.$$

Lemma 2.3.14 is used to get the penultimate equality. Similarly we get two more terms,

$$\lim_{N-M \to \infty} \left(\frac{N-M}{N-M-(K-x)}\right)^{N-M} = e^{K-x}$$

and

$$\lim_{N \to \infty} \left(\frac{N-K}{N}\right)^N = e^{-K}.$$

Note, the product of these three limits is one. Three other terms are

$$\lim M \to \infty \left(\frac{M}{M-x}\right)^{1/2} = 1$$
$$\lim_{N-M\to\infty} \left(\frac{N-M}{N-M-(K-x)}\right)^{1/2} = 1$$

and

$$\lim_{N\to\infty} \left(\frac{N-K}{N}\right)^{1/2} = 1.$$

The only term left is

$$\begin{aligned}
&\lim_{M/N\to p, M\to\infty, N\to\infty} \frac{(M-x)^x (N-M-(K-x))^{K-x}}{(N-K)^K} \\
&= \lim_{M/N\to p, M\to\infty, N\to\infty} \left(\frac{M-x}{N-K}\right)^x \left(\frac{N-M-(K-x)}{N-K}\right)^{K-x} \\
&= p^x(1-p)^{K-x}.
\end{aligned}$$

b. If in (a) we in addition have $K \to \infty$, $p \to 0$, $MK/N \to pK \to \lambda$, by the Poisson approximation to the binomial, we heuristically get

$$\frac{\binom{M}{x}\binom{N-M}{K-x}}{\binom{N}{K}} \to \binom{K}{x} p^x (1-p)^{K-x} \to \frac{e^{-\lambda}\lambda^x}{x!}.$$

c. Using Stirling's formula as in (a), we get

$$\begin{aligned}
&\lim_{N,M,K\to\infty, \frac{M}{N}\to 0, \frac{KM}{N}\to\lambda} \frac{\binom{M}{x}\binom{N-M}{K-x}}{\binom{N}{K}} \\
&= \lim_{N,M,K\to\infty, \frac{M}{N}\to 0, \frac{KM}{N}\to\lambda} \frac{e^{-x}}{x!} \frac{K^x e^x M^x e^x (N-M)^{K-x} e^{K-x}}{N^K e^K} \\
&= \frac{1}{x!} \lim_{N,M,K\to\infty, \frac{M}{N}\to 0, \frac{KM}{N}\to\lambda} \left(\frac{KM}{N}\right)^x \left(\frac{N-M}{N}\right)^{K-x} \\
&= \frac{1}{x!}\lambda^x \lim_{N,M,K\to\infty, \frac{M}{N}\to 0, \frac{KM}{N}\to\lambda} \left(1 - \frac{\frac{MK}{N}}{K}\right)^K \\
&= \frac{e^{-\lambda}\lambda^x}{x!}.
\end{aligned}$$

3.12 Consider a sequence of Bernoulli trials with success probability p. Define X = number of successes in first n trials and Y = number of failures before the rth success. Then X and Y have the specified binomial and hypergeometric distributions, respectively. And we have

$$\begin{aligned}
F_x(r-1) &= P(X \le r-1) \\
&= P(r\text{th success on } (n+1)\text{st or later trial}) \\
&= P(\text{at least } n+1-r \text{ failures before the } r\text{th success}) \\
&= P(Y \ge n-r+1) \\
&= 1 - P(Y \le n-r) \\
&= 1 - F_Y(n-r).
\end{aligned}$$

3.13 For any X with support $0, 1, \ldots$, we have the mean and variance of the $0-$*truncated* X_T are given by

$$\begin{aligned}
\mathrm{E}X_T &= \sum_{x=1}^{\infty} xP(X_T = x) = \sum_{x=1}^{\infty} x\frac{P(X = x)}{P(X > 0)} \\
&= \frac{1}{P(X > 0)}\sum_{x=1}^{\infty} xP(X = x) = \frac{1}{P(X > 0)}\sum_{x=0}^{\infty} xP(X = x) = \frac{\mathrm{E}X}{P(X > 0)}.
\end{aligned}$$

In a similar way we get $\mathrm{E}X_T^2 = \frac{\mathrm{E}X^2}{P(X>0)}$. Thus,

$$\mathrm{Var}X_T = \frac{\mathrm{E}X^2}{P(X > 0)} - \left(\frac{\mathrm{E}X}{P(X > 0)}\right)^2.$$

a. For Poisson(λ), $P(X > 0) = 1 - P(X = 0) = 1 - \frac{e^{-\lambda}\lambda^0}{0!} = 1 - e^{-\lambda}$, therefore

$$\begin{aligned}
P(X_T = x) &= \frac{e^{-\lambda}\lambda^x}{x!(1-e^{-\lambda})} \quad x = 1, 2, \ldots \\
\mathrm{E}X_T &= \lambda/(1 - e^{-\lambda}) \\
\mathrm{Var}X_T &= (\lambda^2 + \lambda)/(1 - e^{-\lambda}) - (\lambda/(1 - e^{-\lambda}))^2.
\end{aligned}$$

b. For negative binomial(r, p), $P(X > 0) = 1 - P(X = 0) = 1 - \binom{r-1}{0}p^r(1-p)^0 = 1 - p^r$. Then

$$\begin{aligned}
P(X_T = x) &= \frac{\binom{r+x-1}{x}p^r(1-p)^x}{1-p^r}, \quad x = 1, 2, \ldots \\
\mathrm{E}X_T &= \frac{r(1-p)}{p(1-p^r)} \\
\mathrm{Var}X_T &= \frac{r(1-p) + r^2(1-p)^2}{p^2(1-p^r)} - \left[\frac{r(1-p)}{p(1-p^r)^2}\right].
\end{aligned}$$

3.14 a. $\sum_{x=1}^{\infty} \frac{-(1-p)^x}{x \log p} = \frac{1}{\log p} \sum_{x=1}^{\infty} \frac{-(1-p)^x}{x} = 1$, since the sum is the Taylor series for $\log p$.

b.

$$\mathrm{E}X = \frac{-1}{\log p}\left[\sum_{x=1}^{\infty}(1-p)^x\right] = \frac{-1}{\log p}\left[\sum_{x=0}^{\infty}(1-p)^x - 1\right] == \frac{-1}{\log p}\left[\frac{1}{p} - 1\right] = \frac{-1}{\log p}\left(\frac{1-p}{p}\right).$$

Since the geometric series converges uniformly,

$$\begin{aligned}
\mathrm{E}X^2 &= \frac{-1}{\log p}\sum_{x=1}^{\infty} x(1-p)^x = \frac{(1-p)}{\log p}\sum_{x=1}^{\infty}\frac{d}{dp}(1-p)^x \\
&= \frac{(1-p)}{\log p}\frac{d}{dp}\sum_{x=1}^{\infty}(1-p)^x = \frac{(1-p)}{\log p}\frac{d}{dp}\left[\frac{1-p}{p}\right] = \frac{-(1-p)}{p^2 \log p}.
\end{aligned}$$

Thus

$$\mathrm{Var}X = \frac{-(1-p)}{p^2 \log p}\left[1 + \frac{(1-p)}{\log p}\right].$$

Alternatively, the mgf can be calculated,

$$M_x(t) = \frac{-1}{\log p}\sum_{x=1}^{\infty}\left[(1-p)e^t\right]^x = \frac{\log(1+pe^t-e^t)}{\log p}$$

and can be differentiated to obtain the moments.

3.15 The moment generating function for the negative binomial is

$$M(t)=\left(\frac{p}{1-(1-p)e^t}\right)^r=\left(1+\frac{1}{r}\frac{r(1-p)(e^t-1)}{1-(1-p)e^t}\right)^r,$$

the term

$$\frac{r(1-p)(e^t-1)}{1-(1-p)e^t}\to\frac{\lambda(e^t-1)}{1}=\lambda(e^t-1)\quad\text{as } r\to\infty,\ p\to 1 \text{ and } r(p-1)\to\lambda.$$

Thus by Lemma 2.3.14, the negative binomial moment generating function converges to $e^{\lambda(e^t-1)}$, the Poisson moment generating function.

3.16 a. Using integration by parts with, $u=t^\alpha$ and $dv=e^{-t}dt$, we obtain

$$\Gamma(\alpha+1)=\int_0^\infty t^{(\alpha+1)-1}e^{-t}dt=t^\alpha(-e^{-t})\Big|_0^\infty-\int_0^\infty \alpha t^{\alpha-1}(-e^{-t})dt=0+\alpha\Gamma(\alpha)=\alpha\Gamma(\alpha).$$

b. Making the change of variable $z=\sqrt{2t}$, i.e., $t=z^2/2$, we obtain

$$\Gamma(1/2)=\int_0^\infty t^{-1/2}e^{-t}dt=\int_0^\infty\frac{\sqrt{2}}{z}e^{-z^2/2}zdz=\sqrt{2}\int_0^\infty e^{-z^2/2}dz=\sqrt{2}\frac{\sqrt{\pi}}{\sqrt{2}}=\sqrt{\pi}.$$

where the penultimate equality uses (3.3.14).

3.17

$$\begin{aligned}\mathrm{E}X^\nu &= \int_0^\infty x^\nu\frac{1}{\Gamma(\alpha)\beta^\alpha}x^{\alpha-1}e^{-x/\beta}dx = \frac{1}{\Gamma(\alpha)\beta^\alpha}\int_0^\infty x^{(\nu+\alpha)-1}e^{-x/\beta}dx\\ &= \frac{\Gamma(\nu+\alpha)\beta^{\nu+\alpha}}{\Gamma(\alpha)\beta^\alpha} = \frac{\beta^\nu\Gamma(\nu+\alpha)}{\Gamma(\alpha)}.\end{aligned}$$

Note, this formula is valid for all $\nu>-\alpha$. The expectation does not exist for $\nu\le-\alpha$.

3.18 If $Y\sim$ negative binomial(r,p), its moment generating function is $M_Y(t)=\left(\frac{p}{1-(1-p)e^t}\right)^r$, and, from Theorem 2.3.15, $M_{pY}(t)=\left(\frac{p}{1-(1-p)e^{pt}}\right)^r$. Now use L'Hôpital's rule to calculate

$$\lim_{p\to 0}\left(\frac{p}{1-(1-p)e^{pt}}\right)=\lim_{p\to 0}\frac{1}{(p-1)te^{pt}+e^{pt}}=\frac{1}{1-t},$$

so the moment generating function converges to $(1-t)^{-r}$, the moment generating function of a gamma$(r,1)$.

3.19 Repeatedly apply the integration-by-parts formula

$$\frac{1}{\Gamma(n)}\int_x^\infty z^{n-1}z^{-z}dz=\frac{x^{n-1}e^{-x}}{(n-1)!}+\frac{1}{\Gamma(n-1)}\int_x^\infty z^{n-2}z^{-z}dz,$$

until the exponent on the second integral is zero. This will establish the formula. If $X\sim$ gamma$(\alpha,1)$ and $Y\sim$ Poisson(x). The probabilistic relationship is $P(X\ge x)=P(Y\le\alpha-1)$.

3.21 The moment generating function would be defined by $\frac{1}{\pi}\int_{-\infty}^\infty\frac{e^{tx}}{1+x^2}dx$. On $(0,\infty)$, $e^{tx}>x$, hence

$$\int_0^\infty\frac{e^{tx}}{1+x^2}dx>\int_0^\infty\frac{x}{1+x^2}dx=\infty,$$

thus the moment generating function does not exist.

3.22 a.

$$\begin{aligned}
\mathrm{E}(X(X-1)) &= \sum_{x=0}^{\infty} x(x-1)\frac{e^{-\lambda}\lambda^x}{x!} \\
&= e^{-\lambda}\lambda^2 \sum_{x=2}^{\infty} \frac{\lambda^{x-2}}{(x-2)!} \quad (\text{let } y=x-2) \\
&= e^{-\lambda}\lambda^2 \sum_{y=0}^{\infty} \frac{\lambda^y}{y!} \;=\; e^{-\lambda}\lambda^2 e^{\lambda} \;=\; \lambda^2 \\
\mathrm{E}X^2 &= \lambda^2 + \mathrm{E}X \;=\; \lambda^2+\lambda \\
\mathrm{Var}X &= \mathrm{E}X^2 - (\mathrm{E}X)^2 \;=\; \lambda^2+\lambda-\lambda^2 \;=\; \lambda.
\end{aligned}$$

b.

$$\begin{aligned}
\mathrm{E}(X(X-1)) &= \sum_{x=0}^{\infty} x(x-1)\binom{r+x-1}{x} pr(1-p)^x \\
&= \sum_{x=2}^{\infty} r(r+1)\binom{r+x-1}{x-2} pr(1-p)^x \\
&= r(r+1)\frac{(1-p)^2}{p^2}\sum_{y=0}^{\infty}\binom{r+2+y-1}{y} pr+2(1-p)^y \\
&= r(r-1)\frac{(1-p)^2}{p^2},
\end{aligned}$$

where in the second equality we substituted $y = x-2$, and in the third equality we use the fact that we are summing over a negative binomial$(r+2,p)$ pmf. Thus,

$$\begin{aligned}
\mathrm{Var}X &= \mathrm{E}X(X-1) + \mathrm{E}X - (\mathrm{E}X)^2 \\
&= r(r+1)\frac{(1-p)^2}{p^2} + \frac{r(1-p)}{p} - \frac{r^2(1-p)^2}{p^2} \\
&= \frac{r(1-p)}{p^2}.
\end{aligned}$$

c.

$$\begin{aligned}
\mathrm{E}X^2 &= \int_0^{\infty} x^2 \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1}e^{-x/\beta}dx \;=\; \frac{1}{\Gamma(\alpha)\beta^\alpha}\int_0^{\infty} x^{\alpha+1}e^{-x/\beta}dx \\
&= \frac{1}{\Gamma(\alpha)\beta^\alpha}\Gamma(\alpha+2)\beta^{\alpha+2} \;=\; \alpha(\alpha+1)\beta^2. \\
\mathrm{Var}X &= \mathrm{E}X^2 - (\mathrm{E}X)^2 \;=\; \alpha(\alpha+1)\beta^2 - \alpha^2\beta^2 \;=\; \alpha\beta^2.
\end{aligned}$$

d. (Use 3.3.18)

$$\begin{aligned}
\mathrm{E}X &= \frac{\Gamma(\alpha+1)\Gamma(\alpha+\beta)}{\Gamma(\alpha+\beta+1)\Gamma(\alpha)} \;=\; \frac{\alpha\Gamma(\alpha)\Gamma(\alpha+\beta)}{(\alpha+\beta)\Gamma(\alpha+\beta)\Gamma(\alpha)} \;=\; \frac{\alpha}{\alpha+\beta}. \\
\mathrm{E}X^2 &= \frac{\Gamma(\alpha+2)\Gamma(\alpha+\beta)}{\Gamma(\alpha+\beta+2)\Gamma(\alpha)} \;=\; \frac{(\alpha+1)\alpha\Gamma(\alpha)\Gamma(\alpha+\beta)}{(\alpha+\beta+1)(\alpha+\beta)\Gamma(\alpha+\beta)\Gamma(\alpha)} \;=\; \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)}. \\
\mathrm{Var}X &= \mathrm{E}X^2 - (\mathrm{E}X)^2 \;=\; \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)} - \frac{\alpha^2}{(\alpha+\beta)^2} \;=\; \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}.
\end{aligned}$$

e. The double exponential(μ,σ) pdf is symmetric about μ. Thus, by Exercise 2.26, E$X=\mu$.

$$\begin{aligned}\mathrm{Var}X &= \int_{-\infty}^{\infty}(x-\mu)^2\frac{1}{2\sigma}e^{-|x-\mu|/\sigma}dx = \int_{-\infty}^{\infty}\sigma z^2\frac{1}{2}e^{-|z|}\sigma dz \\ &= \sigma^2\int_0^{\infty}z^2e^{-z}dz = \sigma^2\Gamma(3) = 2\sigma^2.\end{aligned}$$

3.23 a.

$$\int_{\alpha}^{\infty}x^{-\beta-1}dx = \frac{-1}{\beta}x^{-\beta}\Big|_{\alpha}^{\infty} = \frac{1}{\beta\alpha^{\beta}},$$

thus $f(x)$ integrates to 1 .

b. E$X^n = \frac{\beta\alpha^n}{(n-\beta)}$, therefore

$$\begin{aligned}\mathrm{E}X &= \frac{\alpha\beta}{(1-\beta)}\\ \mathrm{E}X^2 &= \frac{\alpha\beta^2}{(2-\beta)}\\ \mathrm{Var}X &= \frac{\alpha\beta^2}{2-\beta}-\frac{(\alpha\beta)^2}{(1-\beta)^2}\end{aligned}$$

c. If $\beta<2$ the integral of the second moment is infinite.

3.24 a. $f_x(x)=\frac{1}{\beta}e^{-x/\beta}$, $x>0$. For $Y=X^{1/\gamma}$, $f_Y(y)=\frac{\gamma}{\beta}e^{-y^{\gamma}/\beta}y^{\gamma-1}$, $y>0$. Using the transformation $z=y^{\gamma}/\beta$, we calculate

$$\mathrm{E}Y^n = \frac{\gamma}{\beta}\int_0^{\infty}y^{\gamma+n-1}e^{-y^{\gamma}/\beta}dy = \beta^{n/\gamma}\int_0^{\infty}z^{n/\gamma}e^{-z}dz = \beta^{n/\gamma}\Gamma\left(\frac{n}{\gamma}+1\right).$$

Thus E$Y=\beta^{1/\gamma}\Gamma(\frac{1}{\gamma}+1)$ and Var$Y=\beta^{2/\gamma}\left[\Gamma\left(\frac{2}{\gamma}+1\right)-\Gamma^2\left(\frac{1}{\gamma}+1\right)\right]$.

b. $f_x(x)=\frac{1}{\beta}e^{-x/\beta}$, $x>0$. For $Y=(2X/\beta)^{1/2}$, $f_Y(y)=ye^{-y^2/2}$, $y>0$. We now notice that

$$\mathrm{E}Y=\int_0^{\infty}y^2e^{-y^2/2}dy=\frac{\sqrt{2\pi}}{2}$$

since $\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}y^2e^{-y^2/2}=1$, the variance of a standard normal, and the integrand is symmetric. Use integration-by-parts to calculate the second moment

$$\mathrm{E}Y^2=\int_0^{\infty}y^3e^{-y^2/2}dy=2\int_0^{\infty}ye^{-y^2/2}dy=2,$$

where we take $u=y^2$, $dv=ye^{-y^2/2}$. Thus Var$Y=2(1-\pi/4)$.

c. The gamma(a,b) density is

$$f_X(x)=\frac{1}{\Gamma(a)b^a}x^{a-1}e^{-x/b}.$$

Make the transformation $y=1/x$ with $dx=-dy/y^2$ to get

$$f_Y(y)=f_X(1/y)|1/y^2|=\frac{1}{\Gamma(a)b^a}\left(\frac{1}{y}\right)^{a+1}e^{-1/by}.$$

The first two moments are

$$\begin{aligned} \mathrm{E}Y &= \frac{1}{\Gamma(a)b^a}\int_0^\infty \left(\frac{1}{y}\right)^a e^{-1/by} = \frac{\Gamma(a-1)b^{a-1}}{\Gamma(a)b^a} = \frac{1}{(a-1)b} \\ \mathrm{E}Y^2 &= \frac{\Gamma(a-2)b^{a-2}}{\Gamma(a)b^a} = \frac{1}{(a-1)(a-2)b^2}, \end{aligned}$$

and so $\mathrm{Var}Y = \frac{1}{(a-1)^2(a-2)b^2}$.

d. $f_x(x) = \frac{1}{\Gamma(3/2)\beta^{3/2}}x^{3/2-1}e^{-x/\beta}$, $x>0$. For $Y=(X/\beta)^{1/2}$, $f_Y(y) = \frac{2}{\Gamma(3/2)}y^2e^{-y^2}$, $y>0$. To calculate the moments we use integration-by-parts with $u=y^2$, $dv = ye^{-y^2}$ to obtain

$$\mathrm{E}Y = \frac{2}{\Gamma(3/2)}\int_0^\infty y^3e^{-y^2}dy = \frac{2}{\Gamma(3/2)}\int_0^\infty ye^{-y^2}dy = \frac{1}{\Gamma(3/2)}$$

and with $u=y^3, dv=ye^{-y^2}$ to obtain

$$\mathrm{E}Y^2 = \frac{2}{\Gamma(3/2)}\int_0^\infty y^4e^{-y^2}dy = \frac{3}{\Gamma(3/2)}\int_0^\infty y^2e^{-y^2}dy = \frac{3}{\Gamma(3/2)}\sqrt{\pi}.$$

Using the fact that $\frac{1}{2\sqrt{\pi}}\int_{-\infty}^{\infty} y^2e^{-y^2} = 1$, since it is the variance of a $\mathrm{n}(0,2)$, symmetry yields $\int_0^\infty y^2e^{-y^2}dy = \sqrt{\pi}$. Thus, $\mathrm{Var}Y = 6-4/\pi$, using $\Gamma(3/2)=\frac{1}{2}\sqrt{\pi}$.

e. $f_x(x) = e^{-x}$, $x>0$. For $Y=\alpha-\gamma\log X$, $f_Y(y) = e^{-e^{\frac{\alpha-y}{\gamma}}}e^{\frac{\alpha-y}{\gamma}}\frac{1}{\gamma}$, $-\infty<y<\infty$. Calculation of $\mathrm{E}Y$ and $\mathrm{E}Y^2$ cannot be done in closed form. If we define

$$I_1 = \int_0^\infty \log x e^{-x}dx, \qquad I_2 = \int_0^\infty (\log x)^2e^{-x}dx,$$

then $\mathrm{E}Y = \mathrm{E}(\alpha-\gamma\log x) = \alpha-\gamma I_1$, and $\mathrm{E}Y^2 = \mathrm{E}(\alpha-\gamma\log x)^2 = \alpha^2-2\alpha\gamma I_1+\gamma^2 I_2$.The constant $I_1 = .5772157$ is called Euler's constant.

3.25 Note that if T is continuous then,

$$\begin{aligned} P(t\le T\le t+\delta|t\le T) &= \frac{P(t\le T\le t+\delta, t\le T)}{P(t\le T)} \\ &= \frac{P(t\le T\le t+\delta)}{P(t\le T)} \\ &= \frac{F_T(t+\delta)-F_T(t)}{1-F_T(t)}. \end{aligned}$$

Therefore from the definition of derivative,

$$h_T(t) = \frac{1}{1-F_T(t)} = \lim_{\delta\to 0}\frac{F_T(t+\delta)-F_T(t)}{\delta} = \frac{F_T'(t)}{1-F_T(t)} = \frac{f_T(t)}{1-F_T(t)}.$$

Also,

$$-\frac{d}{dt}(\log[1-F_T(t)]) = -\frac{1}{1-F_T(t)}(-f_T(t)) = h_T(t).$$

3.26 a. $f_T(t) = \frac{1}{\beta}e^{-t/\beta}$ and $F_T(t) = \int_0^t \frac{1}{\beta}e^{-x/\beta}dx = -\left.e^{-x/\beta}\right|_0^t = 1-e^{-t/\beta}$. Thus,

$$h_T(t) = \frac{f_T(t)}{1-F_T(t)} = \frac{(1/\beta)e^{-t/\beta}}{1-(1-e^{-t/\beta})} = \frac{1}{\beta}.$$

b. $f_T(t) = \frac{\gamma}{\beta} t^{\gamma-1} e^{-t^\gamma/\beta}, t \geq 0$ and $F_T(t) = \int_0^t \frac{\gamma}{\beta} x^{\gamma-1} e^{-x^\gamma/\beta} dx = \int_0^{t^{\gamma/\beta}} e^{-u} du = -\left. e^{-u}\right|_0^{t^{\gamma/\beta}} = 1 - e^{-t^\gamma/\beta}$, where $u = x^{\gamma/\beta}$. Thus,

$$h_T(t) = \frac{(\gamma/\beta)t^{\gamma-1} e^{-t^\gamma/\beta}}{e^{-t^\gamma/\beta}} = \frac{\gamma}{\beta} t^{\gamma-1}.$$

c. $F_T(t) = \frac{1}{1+e^{-(t-\mu)/\beta}}$ and $f_T(t) = \frac{e^{-(t-\mu)/\beta}}{\left(1+e^{-(t-\mu)/\beta}\right)^2}$. Thus,

$$h_T(t) = \frac{1}{\beta} e^{-(t-\mu)/\beta \left(1+e^{-(t-\mu)/\beta}\right)^2} \frac{1}{\frac{e^{-(t-\mu)/\beta}}{1+e^{-(t-\mu)/\beta}}} = \frac{1}{\beta} F_T(t).$$

3.27 a. The uniform pdf satisfies the inequalities of Exercise 2.27, hence is unimodal.

b. For the gamma(α,β) pdf $f(x)$, ignoring constants, $\frac{d}{dx} f(x) = \frac{x^{\alpha-2} e^{-x/\beta}}{\beta} [\beta(\alpha-1) - x]$, which only has one sign change. Hence the pdf is unimodal with mode $\beta(\alpha-1)$.

c. For the n(μ,σ^2) pdf $f(x)$, ignoring constants, $\frac{d}{dx} f(x) = \frac{x-\mu}{\sigma^2} e^{-(-x/\beta)^2/2\sigma^2}$, which only has one sign change. Hence the pdf is unimodal with mode μ.

d. For the beta(α,β) pdf $f(x)$, ignoring constants,

$$\frac{d}{dx} f(x) = x^{\alpha-2}(1-x)^{\beta-2} \left[(\alpha-1) - x(\alpha+\beta-2)\right],$$

which only has one sign change. Hence the pdf is unimodal with mode $\frac{\alpha-1}{\alpha+\beta-2}$.

3.28 a. (i) μ known,

$$f(x|\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-1}{2\sigma^2}(x-\mu)^2\right),$$

$h(x) = 1, \quad c(\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} I_{(0,\infty)}(\sigma^2), \quad w_1(\sigma^2) = -\frac{1}{2\sigma^2}, \quad t_1(x) = (x-\mu)^2.$

(ii) σ^2 known,

$$f(x|\mu) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right) \exp\left(-\frac{\mu^2}{2\sigma^2}\right) \exp\left(\mu\frac{x}{\sigma^2}\right),$$

$h(x) = \exp\left(\frac{-x^2}{2\sigma^2}\right), \quad c(\mu) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-\mu^2}{2\sigma^2}\right), \quad w_1(\mu) = \mu, \quad t_1(x) = \frac{x}{\sigma^2}.$

b. (i) α known,

$$f(x|\beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{\frac{-x}{\beta}},$$

$h(x) = \frac{x^{\alpha-1}}{\Gamma(\alpha)}, \ x > 0, \quad c(\beta) = \frac{1}{\beta^\alpha}, \quad w_1(\beta) = \frac{1}{\beta}, \quad t_1(x) = -x.$

(ii) β known,

$$f(x|\alpha) = e^{-x/\beta} \frac{1}{\Gamma(\alpha)\beta^\alpha} \exp((\alpha-1)\log x),$$

$h(x) = e^{-x/\beta}, \ x > 0, \quad c(\alpha) = \frac{1}{\Gamma(\alpha)\beta^\alpha} \quad w_1(\alpha) = \alpha - 1, \quad t_1(x) = \log x.$

(iii) α, β unknown,

$$f(x|\alpha,\beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} \exp((\alpha-1)\log x - \frac{x}{\beta}),$$

$h(x) = I_{\{x>0\}}(x), \quad c(\alpha,\beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha}, \quad w_1(\alpha) = \alpha - 1, \quad t_1(x) = \log x,$
$w_2(\alpha,\beta) = -1/\beta, \quad t_2(x) = x.$

c. (i) α known, $h(x) = x^{\alpha-1} I_{[0,1]}(x), \quad c(\beta) = \frac{1}{B(\alpha,\beta)}, \quad w_1(\beta) = \beta - 1, \quad t_1(x) = \log(1-x).$

(ii) β known, $h(x) = (1-x)^{\beta-1} I_{[0,1]}(x), \quad c(\alpha) = \frac{1}{B(\alpha,\beta)}, \quad w_1(x) = \alpha - 1, \quad t_1(x) = \log x.$

(iii) α, β unknown,
$h(x) = I_{[0,1]}(x)$, $c(\alpha,\beta) = \frac{1}{B(\alpha,\beta)}$, $w_1(\alpha) = \alpha - 1$, $t_1(x) = \log x$,
$w_2(\beta) = \beta - 1$, $t_2(x) = \log(1-x)$.

d. $h(x) = \frac{1}{x!} I_{\{0,1,2,\ldots\}}(x)$, $c(\theta) = e^{-\theta}$, $w_1(\theta) = \log\theta$, $t_1(x) = x$.

e. $h(x) = \binom{x-1}{r-1} I_{\{r,r+1,\ldots\}}(x)$, $c(p) = \left(\frac{p}{1-p}\right)^r$, $w_1(p) = \log(1-p)$, $t_1(x) = x$.

3.29 a. For the n(μ,σ^2)

$$f(x) = \left(\frac{1}{\sqrt{2\pi}}\right)\left(\frac{e^{-\mu^2/2\sigma^2}}{\sigma}\right)\left(e^{-x^2/2\sigma^2 + x\mu/\sigma^2}\right),$$

so the natural parameter is $(\eta_1,\eta_2) = (-1/2\sigma^2, \mu/\sigma^2)$ with natural parameter space $\{(\eta_1,\eta_2):\eta_1 < 0, -\infty < \eta_2 < \infty\}$.

b. For the gamma(α,β),

$$f(x) = \left(\frac{1}{\Gamma(\alpha)\beta^\alpha}\right)\left(e^{(\alpha-1)\log x - x/\beta}\right),$$

so the natural parameter is $(\eta_1,\eta_2) = (\alpha-1, -1/\beta)$ with natural parameter space $\{(\eta_1,\eta_2):\eta_1 > -1, \eta_2 < 0\}$.

c. For the beta(α,β),

$$f(x) = \left(\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\right)\left(e^{(\alpha-1)\log x + (\beta-1)\log(1-x)}\right),$$

so the natural parameter is $(\eta_1,\eta_2) = (\alpha-1, \beta-1)$ and the natural parameter space is $\{(\eta_1,\eta_2):\eta_1 > -1, \eta_2 > -1\}$.

d. For the Poisson

$$f(x) = \left(\frac{1}{x!}\right)\left(e^{-\theta}\right)e^{x\log\theta}$$

so the natural parameter is $\eta = \log\theta$ and the natural parameter space is $\{\eta: -\infty < \eta < \infty\}$.

e. For the negative binomial(r,p), r known,

$$P(X=x) = \binom{r+x-1}{x}(p^r)\left(e^{x\log(1-p)}\right),$$

so the natural parameter is $\eta = \log(1-p)$ with natural parameter space $\{\eta:\eta < 0\}$.

3.31 a.

$$\begin{aligned}
0 &= \frac{\partial}{\partial\theta}\int h(x)c(\theta)\exp\left(\sum_{i=1}^k w_i(\theta)t_i(x)\right)dx \\
&= \int h(x)c'(\theta)\exp\left(\sum_{i=1}^k w_i(\theta)t_i(x)\right)dx \\
&\quad + \int h(x)c(\theta)\exp\left(\sum_{i=1}^k w_i(\theta)t_i(x)\right)\left(\sum_{i=1}^k \frac{\partial w_i(\theta)}{\partial\theta_j}t_i(x)\right)dx \\
&= \int h(x)\left[\frac{\partial}{\partial\theta_j}\log c(\theta)\right]c(\theta)\exp\left(\sum_{i=1}^k w_i(\theta)t_i(x)\right)dx + \mathrm{E}\left[\sum_{i=1}^k \frac{\partial w_i(\theta)}{\partial\theta_j}t_i(x)\right] \\
&= \frac{\partial}{\partial\theta_j}\log c(\theta) + \mathrm{E}\left[\sum_{i=1}^k \frac{\partial w_i(\theta)}{\partial\theta_j}t_i(x)\right]
\end{aligned}$$

Therefore $\mathrm{E}\left[\sum_{i=1}^k \frac{\partial w_i(\theta)}{\partial\theta_j}t_i(x)\right] = -\frac{\partial}{\partial\theta_j}\log c(\theta)$.

b.

$$
\begin{aligned}
0 &= \frac{\partial^2}{\partial\theta^2}\int h(x)c(\theta)\exp\left(\sum_{i=1}^k w_i(\theta)t_i(x)\right)dx \\
&= \int h(x)c''(\theta)\exp\left(\sum_{i=1}^k w_i(\theta)t_i(x)\right)dx \\
&\quad +\int h(x)c'(\theta)\exp\left(\sum_{i=1}^k w_i(\theta)t_i(x)\right)\left(\sum_{i=1}^k \frac{\partial w_i(\theta)}{\partial\theta_j}t_i(x)\right)dx \\
&\quad +\int h(x)c'(\theta)\exp\left(\sum_{i=1}^k w_i(\theta)t_i(x)\right)\left(\sum_{i=1}^k \frac{\partial w_i(\theta)}{\partial\theta_j}t_i(x)\right)dx \\
&\quad +\int h(x)c(\theta)\exp\left(\sum_{i=1}^k w_i(\theta)t_i(x)\right)\left(\sum_{i=1}^k \frac{\partial w_i(\theta)}{\partial\theta_j}t_i(x)\right)^2dx \\
&\quad +\int h(x)c(\theta)\exp\left(\sum_{i=1}^k w_i(\theta)t_i(x)\right)\left(\sum_{i=1}^k \frac{\partial^2 w_i(\theta)}{\partial\theta_j^2}t_i(x)\right)dx \\
&= \int h(x)\left[\frac{\partial^2}{\partial\theta_j^2}\log c(\theta)\right]c(\theta)\exp\left(\sum_{i=1}^k w_i(\theta)t_i(x)\right)dx \\
&\quad +\int h(x)\left[\frac{c'(\theta)}{c(\theta)}\right]^2 c(\theta)\exp\left(\sum_{i=1}^k w_i(\theta)t_i(x)\right)dx \\
&\quad +2\left(\frac{\partial}{\partial\theta_j}\log c(\theta)\right)\mathrm{E}\left[\sum_{i=1}^k \frac{\partial w_i(\theta)}{\partial\theta_j}t_i(x)\right] \\
&\quad +\mathrm{E}\left[(\sum_{i=1}^k \frac{\partial w_i(\theta)}{\partial\theta_j}t_i(x))^2\right]+\mathrm{E}\left[\sum_{i=1}^k \frac{\partial^2 w_i(\theta)}{\partial\theta_j^2}t_i(x)\right] \\
&= \frac{\partial^2}{\partial\theta_j^2}\log c(\theta)+\left[\frac{\partial}{\partial\theta_j}\log c(\theta)\right]^2 \\
&\quad -2\mathrm{E}\left[\sum_{i=1}^k \frac{\partial w_i(\theta)}{\partial\theta_j}t_i(x)\right]\mathrm{E}\left[\sum_{i=1}^k \frac{\partial w_i(\theta)}{\partial\theta_j}t_i(x)\right] \\
&\quad +\mathrm{E}\left[(\sum_{i=1}^k \frac{\partial w_i(\theta)}{\partial\theta_j}t_i(x))^2\right]+\mathrm{E}\left[\sum_{i=1}^k \frac{\partial^2 w_i(\theta)}{\partial\theta_j^2}t_i(x)\right] \\
&= \frac{\partial^2}{\partial\theta_j^2}\log c(\theta)+\mathrm{Var}\left(\sum_{i=1}^k \frac{\partial w_i(\theta)}{\partial\theta_j}t_i(x)\right)+\mathrm{E}\left[\sum_{i=1}^k \frac{\partial^2 w_i(\theta)}{\partial\theta_j^2}t_i(x)\right].
\end{aligned}
$$

Therefore $\mathrm{Var}\left(\sum_{i=1}^k \frac{\partial w_i(\theta)}{\partial\theta_j}t_i(x)\right) = -\frac{\partial^2}{\partial\theta_j^2}\log c(\theta) - \mathrm{E}\left[\sum_{i=1}^k \frac{\partial^2 w_i(\theta)}{\partial\theta_j^2}t_i(x)\right]$.

3.33 a. (i) $h(x) = e^x I_{\{-\infty<x<\infty\}}(x)$, $c(\theta) = \frac{1}{\sqrt{2\pi\theta}}\exp(\frac{-\theta}{2})\theta > 0$, $w_1(\theta) = \frac{1}{2\theta}$, $t_1(x) = -x^2$.

(ii) The nonnegative real line.

b. (i) $h(x) = I_{\{-\infty<x<\infty\}}(x)$, $c(\theta) = \frac{1}{\sqrt{2\pi a\theta^2}}\exp(\frac{-1}{2a}) - \infty < \theta < \infty, a > 0$,
$w_1(\theta) = \frac{1}{2a\theta^2}$, $w_2(\theta) = \frac{1}{a\theta}$, $t_1(x) = -x^2$, $t_2(x) = x$.

(ii) A parabola.

c. (i) $h(x) = \frac{1}{x} I_{\{0<x<\infty\}}(x)$, $c(\alpha) = \frac{\alpha^\alpha}{\Gamma(\alpha)} \alpha > 0$, $w_1(\alpha) = \alpha$, $w_2(\alpha) = \alpha$, $t_1(x) = \log(x)$, $t_2(x) = -x$.

(ii) A line.

d. (i) $h(x) = C\exp(x^4) I_{\{-\infty<x<\infty\}}(x)$, $c(\theta) = \exp(\theta^4) - \infty < \theta < \infty$, $w_1(\theta) = \theta$, $w_2(\theta) = \theta^2$, $w_3(\theta) = \theta^3$, $t_1(x) = -4x^3$, $t_2(x) = 6x^2$, $t_3(x) = -4x$.

(ii) The curve is a spiral in 3-space.

(iii) A good picture can be generated with the Mathematica statement

```
ParametricPlot3D[{t, t^2, t^3}, {t, 0, 1}, ViewPoint -> {1, -2, 2.5}].
```

3.35 a. In Exercise 3.34(a) $w_1(\lambda) = \frac{1}{2\lambda}$ and for a $\mathrm{n}(e^\theta, e^\theta)$, $w_1(\theta) = \frac{1}{2e^\theta}$.

b. $\mathrm{E}X = \mu = \alpha\beta$, then $\beta = \frac{\mu}{\alpha}$. Therefore $h(x) = \frac{1}{x} I_{\{0<x<\infty\}}(x)$, $c(\alpha) = \frac{\alpha^\alpha}{\Gamma(\alpha)(\frac{\mu}{\alpha})^\alpha}, \alpha > 0$, $w_1(\alpha) = \alpha$, $w_2(\alpha) = \frac{\alpha}{\mu}$, $t_1(x) = \log(x)$, $t_2(x) = -x$.

c. From (b) then $(\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n) = (\alpha_1, \ldots, \alpha_n, \frac{\alpha_1}{\mu}, \ldots, \frac{\alpha_n}{\mu})$

3.37 The pdf $(\frac{1}{\sigma}) f(\frac{(x-\mu)}{\sigma})$ is symmetric about μ because, for any $\epsilon > 0$,

$$\frac{1}{\sigma} f\left(\frac{(\mu+\epsilon)-\mu}{\sigma}\right) = \frac{1}{\sigma} f\left(\frac{\epsilon}{\sigma}\right) = \frac{1}{\sigma} f\left(-\frac{\epsilon}{\sigma}\right) = \frac{1}{\sigma} f\left(\frac{(\mu-\epsilon)-\mu}{\sigma}\right).$$

Thus, by Exercise 2.26b, μ is the median.

3.38 $P(X > x_\alpha) = P(\sigma Z + \mu > \sigma z_\alpha + \mu) = P(Z > z_\alpha)$ by Theorem 3.5.6.

3.39 First take $\mu = 0$ and $\sigma = 1$.

a. The pdf is symmetric about 0, so 0 must be the median. Verifying this, write

$$P(Z \geq 0) = \int_0^\infty \frac{1}{\pi} \frac{1}{1+z^2} dz = \frac{1}{\pi} \tan^{-1}(z) \Big|_0^\infty = \frac{1}{\pi} \left(\frac{\pi}{2} - 0\right) = \frac{1}{2}.$$

b. $P(Z \geq 1) = \frac{1}{\pi} \tan^{-1}(z) \Big|_1^\infty = \frac{1}{\pi} \left(\frac{\pi}{2} - \frac{\pi}{4}\right) = \frac{1}{4}$. By symmetry this is also equal to $P(Z \leq -1)$. Writing $z = (x - \mu)/\sigma$ establishes $P(X \geq \mu) = \frac{1}{2}$ and $P(X \geq \mu + \sigma) = \frac{1}{4}$.

3.40 Let $X \sim f(x)$ have mean μ and variance σ^2. Let $Z = \frac{X-\mu}{\sigma}$. Then

$$\mathrm{E}Z = \left(\frac{1}{\sigma}\right) \mathrm{E}(X - \mu) = 0$$

and

$$\mathrm{Var}Z = \mathrm{Var}\left(\frac{X - \mu}{\sigma}\right) = \left(\frac{1}{\sigma^2}\right) \mathrm{Var}(X - \mu) = \left(\frac{1}{\sigma^2}\right) \mathrm{Var}X = \frac{\sigma^2}{\sigma^2} = 1.$$

Then compute the pdf of Z, $f_Z(z) = f_x(\sigma z + \mu) \cdot \sigma = \sigma f_x(\sigma z + \mu)$ and use $f_Z(z)$ as the standard pdf.

3.41 a. This is a special case of Exercise 3.42a.

b. This is a special case of Exercise 3.42b.

3.42 a. Let $\theta_1 > \theta_2$. Let $X_1 \sim f(x - \theta_1)$ and $X_2 \sim f(x - \theta_2)$. Let $F(z)$ be the cdf corresponding to $f(z)$ and let $Z \sim f(z)$.Then

$$\begin{aligned} F(x \mid \theta_1) &= P(X_1 \leq x) = P(Z + \theta_1 \leq x) = P(Z \leq x - \theta_1) = F(x - \theta_1) \\ &\leq F(x - \theta_2) = P(Z \leq x - \theta_2) = P(Z + \theta_2 \leq x) = P(X_2 \leq x) \\ &= F(x \mid \theta_2). \end{aligned}$$

The inequality is because $x - \theta_2 > x - \theta_1$, and F is nondecreasing. To get strict inequality for some x, let $(a, b]$ be an interval of length $\theta_1 - \theta_2$ with $P(a < Z \leq b) = F(b) - F(a) > 0$. Let $x = a + \theta_1$. Then

$$\begin{aligned} F(x \mid \theta_1) &= F(x-\theta_1) = F(a+\theta_1-\theta_1) = F(a) \\ &< F(b) = F(a+\theta_1-\theta_2) = F(x-\theta_2) = F(x \mid \theta_2). \end{aligned}$$

b. Let $\sigma_1 > \sigma_2$. Let $X_1 \sim f(x/\sigma_1)$ and $X_2 \sim f(x/\sigma_2)$. Let $F(z)$ be the cdf corresponding to $f(z)$ and let $Z \sim f(z)$. Then, for $x > 0$,

$$\begin{aligned} F(x \mid \sigma_1) &= P(X_1 \leq x) = P(\sigma_1 Z \leq x) = P(Z \leq x/\sigma_1) = F(x/\sigma_1) \\ &\leq F(x/\sigma_2) = P(Z \leq x/\sigma_2) = P(\sigma_2 Z \leq x) = P(X_2 \leq x) \\ &= F(x \mid \sigma_2). \end{aligned}$$

The inequality is because $x/\sigma_2 > x/\sigma_1$ (because $x > 0$ and $\sigma_1 > \sigma_2 > 0$), and F is nondecreasing. For $x \leq 0$, $F(x \mid \sigma_1) = P(X_1 \leq x) = 0 = P(X_2 \leq x) = F(x \mid \sigma_2)$. To get strict inequality for some x, let $(a, b]$ be an interval such that $a > 0$, $b/a = \sigma_1/\sigma_2$ and $P(a < Z \leq b) = F(b) - F(a) > 0$. Let $x = a\sigma_1$. Then

$$\begin{aligned} F(x \mid \sigma_1) &= F(x/\sigma_1) = F(a\sigma_1/\sigma_1) = F(a) \\ &< F(b) = F(a\sigma_1/\sigma_2) = F(x/\sigma_2) \\ &= F(x \mid \sigma_2). \end{aligned}$$

3.43 a. $F_Y(y|\theta) = 1 - F_X(\frac{1}{y}|\theta)$ $y > 0$, by Theorem 2.1.3. For $\theta_1 > \theta_2$,

$$F_Y(y|\theta_1) = 1 - F_X\left(\frac{1}{y}\middle|\theta_1\right) \leq 1 - F_X\left(\frac{1}{y}\middle|\theta_2\right) = F_Y(y|\theta_2)$$

for all y, since $F_X(x|\theta)$ is stochastically increasing and if $\theta_1 > \theta_2$, $F_X(x|\theta_2) \leq F_X(x|\theta_1)$ for all x. Similarly, $F_Y(y|\theta_1) = 1 - F_X(\frac{1}{y}|\theta_1) < 1 - F_X(\frac{1}{y}|\theta_2) = F_Y(y|\theta_2)$ for some y, since if $\theta_1 > \theta_2$, $F_X(x|\theta_2) < F_X(x|\theta_1)$ for some x. Thus $F_Y(y|\theta)$ is stochastically decreasing in θ.

b. $F_X(x|\theta)$ is stochastically increasing in θ. If $\theta_1 > \theta_2$ and $\theta_1, \theta_2 > 0$ then $\frac{1}{\theta_2} > \frac{1}{\theta_1}$. Therefore $F_X(x|\frac{1}{\theta_1}) \leq F_X(x|\frac{1}{\theta_2})$ for all x and $F_X(x|\frac{1}{\theta_1}) < F_X(x|\frac{1}{\theta_2})$ for some x. Thus $F_X(x|\frac{1}{\theta})$ is stochastically decreasing in θ.

3.44 The function $g(x) = |x|$ is a nonnegative function. So by Chebychev's Inequality,

$$P(|X| \geq b) \leq \mathrm{E}|X|/b.$$

Also, $P(|X| \geq b) = P(X^2 \geq b^2)$. Since $g(x) = x^2$ is also nonnegative, again by Chebychev's Inequality we have

$$P(|X| \geq b) = P(X^2 \geq b^2) \leq \mathrm{E}X^2/b^2.$$

For $X \sim$ exponential(1), $\mathrm{E}|X| = \mathrm{E}X = 1$ and $\mathrm{E}X^2 = \mathrm{Var}X + (\mathrm{E}X)^2 = 2$. For $b = 3$,

$$\mathrm{E}|X|/b = 1/3 > 2/9 = \mathrm{E}X^2/b^2.$$

Thus $\mathrm{E}X^2/b^2$ is a better bound. But for $b = \sqrt{2}$,

$$\mathrm{E}|X|/b = 1/\sqrt{2} < 1 = \mathrm{E}X^2/b^2.$$

Thus $\mathrm{E}|X|/b$ is a better bound.

3.45 a.

$$\begin{aligned} M_X(t) &= \int_{-\infty}^{\infty} e^{tx} f_X(x)dx \geq \int_a^{\infty} e^{tx} f_X(x)dx \\ &\geq e^{ta} \int_a^{\infty} f_X(x)dx = e^{ta} P(X \geq a), \end{aligned}$$

where we use the fact that e^{tx} is increasing in x for $t > 0$.

b.

$$\begin{aligned} M_X(t) &= \int_{-\infty}^{\infty} e^{tx} f_X(x)dx \geq \int_{-\infty}^{a} e^{tx} f_X(x)dx \\ &\geq e^{ta} \int_{-\infty}^{a} f_X(x)dx = e^{ta} P(X \leq a), \end{aligned}$$

where we use the fact that e^{tx} is decreasing in x for $t < 0$.

c. $h(t, x)$ must be nonnegative.

3.46 For $X \sim$ uniform$(0, 1)$, $\mu = \frac{1}{2}$ and $\sigma^2 = \frac{1}{12}$, thus

$$P(|X - \mu| > k\sigma) = 1 - P\left(\frac{1}{2} - \frac{k}{\sqrt{12}} \leq X \leq \frac{1}{2} + \frac{k}{\sqrt{12}}\right) = \begin{cases} 1 - \frac{2k}{\sqrt{12}} & k < \sqrt{3}, \\ 0 & k \geq \sqrt{3}, \end{cases}$$

For $X \sim$ exponential(λ), $\mu = \lambda$ and $\sigma^2 = \lambda^2$, thus

$$P(|X - \mu| > k\sigma) = 1 - P(\lambda - k\lambda \leq X \leq \lambda + k\lambda) = \begin{cases} 1 + e^{-(k+1)} - e^{k-1} & k \leq 1 \\ e^{-(k+1)} & k > 1. \end{cases}$$

From Example 3.6.2, Chebychev's Inequality gives the bound $P(|X - \mu| > k\sigma) \leq 1/k^2$.

Comparison of probabilities

k	u(0, 1) exact	exp(λ) exact	Chebychev
.1	.942	.926	100
.5	.711	.617	4
1	.423	.135	1
1.5	.134	.0821	.44
$\sqrt{3}$	0	0.0651	.33
2	0	0.0498	.25
4	0	0.00674	.0625
10	0	0.0000167	.01

So we see that Chebychev's Inequality is quite conservative.

3.47

$$\begin{aligned} P(|Z| > t) &= 2P(Z > t) = 2\frac{1}{\sqrt{2\pi}} \int_t^{\infty} e^{-x^2/2} dx \\ &= \sqrt{\frac{2}{\pi}} \int_t^{\infty} \frac{1+x^2}{1+x^2} e^{-x^2/2} dx \\ &= \sqrt{\frac{2}{\pi}} \left[\int_t^{\infty} \frac{1}{1+x^2} e^{-x^2/2} dx + \int_t^{\infty} \frac{x^2}{1+x^2} e^{-x^2/2} dx \right]. \end{aligned}$$

To evaluate the second term, let $u = \frac{x}{1+x^2}$, $dv = xe^{-x^2/2}dx$, $v = -e^{-x^2/2}$, $du = \frac{1-x^2}{(1+x^2)^2}$, to obtain

$$\begin{aligned}\int_t^\infty \frac{x^2}{1+x^2}e^{-x^2/2}dx &= \frac{x}{1+x^2}(-e^{-x^2/2})\Big|_t^\infty - \int_t^\infty \frac{1-x^2}{(1+x^2)^2}(-e^{-x^2/2})dx \\ &= \frac{t}{1+t^2}e^{-t^2/2} + \int_t^\infty \frac{1-x^2}{(1+x^2)^2}e^{-x^2/2}dx.\end{aligned}$$

Therefore,

$$\begin{aligned}P(Z \ge t) &= \sqrt{\frac{2}{\pi}}\frac{t}{1+t^2}e^{-t^2/2} + \sqrt{\frac{2}{\pi}}\int_t^\infty \left(\frac{1}{1+x^2} + \frac{1-x^2}{(1+x^2)^2}\right)e^{-x^2/2}dx \\ &= \sqrt{\frac{2}{\pi}}\frac{t}{1+t^2}e^{-t^2/2} + \sqrt{\frac{2}{\pi}}\int_t^\infty \left(\frac{2}{(1+x^2)^2}\right)e^{-x^2/2}dx \\ &\ge \sqrt{\frac{2}{\pi}}\frac{t}{1+t^2}e^{-t^2/2}.\end{aligned}$$

3.48 For the negative binomial

$$P(X = x+1) = \binom{r+x+1-1}{x+1}p^r(1-p)^{x+1} = \left(\frac{r+x}{x+1}\right)(1-p)P(X=x).$$

For the hypergeometric

$$P(X = x+1) = \begin{cases} \frac{(M-x)(k-x+x+1)(x+1)}{P(X=x)} & \text{if } x < k,\ x < M,\ x \ge M-(N-k) \\ \frac{\binom{M}{x+1}\binom{N-M}{k-x-1}}{\binom{N}{k}} & \text{if } x = M-(N-k)-1 \\ 0 & \text{otherwise.}\end{cases}$$

3.49 a.

$$\mathrm{E}(g(X)(X-\alpha\beta)) = \int_0^\infty g(x)(x-\alpha\beta)\frac{1}{\Gamma(\alpha)}\beta^\alpha x^{\alpha-1}e^{-x/\beta}dx.$$

Let $u = g(x)$, $du = g'(x)$, $dv = (x-\alpha\beta)x^{\alpha-1}e^{-x/\beta}$, $v = -\beta x^\alpha e^{-x/\beta}$. Then

$$\mathrm{E}g(X)(X-\alpha\beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha}\left[-g(x)\beta x^\alpha e^{-x/\beta}\Big|_0^\infty + \beta\int_0^\infty g'(x)x^\alpha e^{-x/\beta}dx\right].$$

Assuming $g(x)$ to be differentiable, $\mathrm{E}|Xg'(X)| < \infty$ and $\lim_{x\to\infty} g(x)x^\alpha e^{-x/\beta} = 0$, the first term is zero, and the second term is $\beta\mathrm{E}(Xg'(X))$.

b.

$$\mathrm{E}\left[g(X)\left(\beta-(\alpha-1)\frac{1-X}{x}\right)\right] = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\int_0^1 g(x)\left(\beta-(\alpha-1)\frac{1-x}{x}\right)x^{\alpha-1}(1-x)^{\beta-1}dx.$$

Let $u = g(x)$ and $dv = (\beta-(\alpha-1)\frac{1-x}{x})x^{\alpha-1}(1-x)^\beta$. The expectation is

$$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\left[g(x)x^{\alpha-1}(1-x)^\beta\Big|_0^1 + \int_0^1 (1-x)g'(x)x^{\alpha-1}(1-x)^{\beta-1}dx\right] = \mathrm{E}((1-X)g'(X)),$$

assuming the first term is zero and the integral exists.

3.50 The proof is similar to that of part a) of Theorem 3.6.8. For $X \sim$ negative binomial(r, p),

$$
\begin{aligned}
&\mathrm{E}g(X)\\
&= \sum_{x=0}^{\infty} g(x)\binom{r+x-1}{x} p^r(1-p)^x\\
&= \sum_{y=1}^{\infty} g(y-1)\binom{r+y-2}{y-1} p^r(1-p)^{y-1} \qquad \text{(set } y = x+1\text{)}\\
&= \sum_{y=1}^{\infty} g(y-1)\left(\frac{y}{r+y-1}\right)\binom{r+y-1}{y} p^r(1-p)^{y-1}\\
&= \sum_{y=0}^{\infty} \left[\frac{y}{r+y-1}\frac{g(y-1)}{1-p}\right]\left[\binom{r+y-1}{y} p^r(1-p)^{y}\right] \qquad \text{(the summand is zero at y = 0)}\\
&= \mathrm{E}\left(\frac{X}{r+X-1}\frac{g(X-1)}{1-p}\right),
\end{aligned}
$$

where in the third equality we use the fact that $\binom{r+y-2}{y-1} = \left(\frac{y}{r+y-1}\right)\binom{r+y-1}{y}$.