1. Estimation in the normal distribution:

   (a) We have
   $$\mathbf{E}(\bar{X}) = \mathbf{E}\{(X_1+\cdots+X_n)/n\} = \{\mathbf{E}(X_1)+\cdots+\mathbf{E}(X_n)\}/n = n\mu/n = \mu.$$

   (b) The symmetry of the distribution of $X_i - \mu$ implies
   $$\mathbf{E}\{(X-\mu)_{(p+1)}\} = \mathbf{E}\{(\mu - X)_{(p+1)}\}.$$
   Combining this with the identity
   $$X_{(p+1)} - \mu = (X-\mu)_{(p+1)} = -\{(\mu - X)_{(p+1)}\}$$
   yields
   $$\mathbf{E}\{X_{(p+1)} - \mu\} = \mathbf{E}\{(X-\mu)_{(p+1)}\} = \mathbf{E}\{(\mu - X)_{(p+1)}\} = -\mathbf{E}\{(\mu - X)_{(p+1)}\} = 0$$
   and hence that $\mathbf{E}(\tilde{\mu}) = \mathbf{E}(X_{(p+1)}) = \mu$.
   Note that the argument applies to any distribution which is symmetric around μ so the unbiasedness does not only hold in the normal case.

   (c) The identity for any positive random variable
   $$\mathbf{V}(\sqrt{Y}) = \mathbf{E}(Y) - \{\mathbf{E}(\sqrt{Y})\}^2$$
   yields for $Y = SSD/n$ where, clearly, $\mathbf{V}(\sqrt{Y}) > 0$
   $$\mathbf{E}(\hat{\sigma}) = \mathbf{E}\{\sqrt{(SSD/n)}\} < \sqrt{\mathbf{E}(SSD/n)} = \sigma,$$
   hence $\hat{\sigma}$ is not unbiased.

   (d) For $n = 3$, $Y = SSD/\sigma^2$ follows a χ^2-distribution with two degrees of freedom, hence
   $$\mathbf{E}(\sqrt{Y}) = \int_0^\infty \sqrt{y}\frac{1}{2\Gamma(1)}e^{-y/2}\,dy = \frac{2^{3/2}\Gamma(3/2)}{2\Gamma(1)} = \sqrt{2}\frac{1}{2}\Gamma(1/2) = \sqrt{\pi/2},$$
   hence $\tilde{\sigma} = 2\hat{\sigma}/\sqrt{\pi}$ is unbiased for σ.

   (e) If we let $d = (n-1)$ we similarly find for general n that
   $$\mathbf{E}(\sqrt{Y}) = \int_0^\infty \sqrt{y}\frac{y^{d/2-1}}{2^{d/2}\Gamma(d/2)}e^{-y/2}\,dy = \frac{2^{(d+1)/2}\Gamma((d+1)/2)}{2^{d/2}\Gamma(d/2)},$$
   so
   $$c = \frac{\sqrt{d}\Gamma(d/2)}{\sqrt{2}\Gamma((d+1)/2)}$$
   makes cS an unbiased estimator of σ.

2. Estimation in the uniform distribution on $(0, \theta)$

   (a) The variance of the average is

$$\mathbf{V}(\tilde{\theta}) = \mathbf{V}(\bar{X}) = \frac{\theta^2}{12n}$$

   as $\mathbf{V}(X_i) = \theta^2/12$, which is seen by direct integration

$$\mathbf{E}(X_i) = \int_0^\theta \frac{1}{\theta} x\,dx = \theta/2, \quad \mathbf{E}(X_i^2) = \theta^2/3, \quad \mathbf{V}(X_i) = \theta^2/3 - \theta^2/4 = \theta^2/12.$$

   (b) The likelihood function is 0 for $0 < \theta < x_{(n)}$ and equal to θ^{-n} for $\theta > x_{(n)}$.

   (c) Let $U = X_{(n)}$, we then have

$$P(U \leq u) = \prod_i^n P(X_i \leq u) = (u/\theta)^n \text{ for } 0 < u < \theta$$

   so differentiation yields that U has density

$$f(u;\theta) = nu^{n-1}\theta^{-n} \text{ for } 0 < u < \theta.$$

   Direct integration now yields

$$\mathbf{E}(\hat{\theta}) = \mathbf{E}(U) = \frac{n\theta}{n+1}$$

   (d) The estimator

$$\check{\theta} = \frac{n+1}{n} X_{(n)}$$

   is unbiased. Direct integration gives $E(U^2) = (n\theta^2)/(n+2)$ so

$$\mathbf{V}(\hat{\theta}) = \mathbf{V}(U) = \frac{n\theta^2}{(n+2)(n+1)^2}$$

   and

$$\mathbf{V}(\check{\theta}) = \frac{\theta^2}{n(n+2)}.$$

   (e) Clearly, the mean square error of $\tilde{\theta}$ is very large compared to the mean square error of $\check{\theta}$. Even the ratio of variances of the second to the first is

$$\frac{12n}{n(n+2)} = \frac{12}{n+2}$$

   which tends to 0 for $n \to \infty$.

The mean square error of $\hat{\theta}$ is

$$\begin{aligned}\mathrm{mse}(\hat{\theta}) &= \mathbf{V}(\hat{\theta}) + \{\mathrm{bias}(\hat{\theta})\}^2 \\ &= \theta^2\left(\frac{n}{(n+2)(n+1)^2} + \frac{1}{(n+1)^2}\right) = \frac{2\theta^2}{(n+2)(n+1)},\end{aligned}$$

so, in terms of mean square error,

$$\mathrm{reff}(\check{\theta} : \hat{\theta}) = \frac{\mathrm{mse}(\hat{\theta})}{\mathrm{mse}(\check{\theta})} = \frac{2n}{n+1}$$

which is larger than 1 unless $n = 1$, so the bias-correction seems well justified in this case.

3. Linear unbiased estimation

(a) We get

$$\mathbf{E}(\hat{\mu}) = \mathbf{E}(w_1X_1 + w_2X_2 + \cdots + w_nX_n) = \sum_i w_i\mathbf{E}(X_i) = \mu\sum_i w_i$$

so $\hat{\mu}$ is unbiased if and only if $\sum w_i = 1$.

(b) We have

$$\mathbf{V}(\hat{\mu}) = \sum_i w_i^2\sigma_i^2,$$

which should be minimized subject to the constraint $\sum_i w_i = 1$. Use now e.g. Lagrange multipliers to get the result:

$$L = \sum_i w_i^2\sigma_i^2 - \lambda(\sum_i w_i - 1),$$

and differentiation w.r.t. w_i and yields

$$2w_i\sigma_i^2 = \lambda \Rightarrow w_i \propto \sigma_i^{-2}$$

so

$$w_i = \frac{\sigma_i^{-2}}{\sum_j \sigma_j^{-2}}.$$

(c) We have

$$\mathbf{V}(\hat{\mu}) = \sum_i w_i^2\sigma_i^2 = \frac{\sum_i \sigma_i^{-4}\sigma_i^2}{\left(\sum \sigma_i^{-2}\right)^2} = \frac{1}{\sum_i \sigma_i^{-2}}.$$

(d) When $\sigma_i^2 = \sigma^2$ we have $\mathbf{V}(\hat{\mu}) = \sigma^2/n$ which tends to zero for $n \to \infty$ whereas

$$\mathrm{bias}(\hat{\mu}) = \sum \beta_i/n = \bar{\beta}$$

is equal to the average bias and

$$\mathrm{mse}(\hat{\mu}) = \sigma^2/n + \bar{\beta}^2.$$

Therefore the bias tends to dominate the variance as n gets large, which is very unfortunate.

(e) The last question is really unfair (sorry about that) and there is plenty of room for imagination. But if, for example, σ_i^2 is bounded, similar phenomena as above will prevail.

4. Estimation in the Gamma distribution

   (a)
$$c(\theta)^{-1} = \int_0^\infty x^2 e^{-\theta x}\,dx = \Gamma(3)/\theta^3 = 2/\theta^3.$$

   (b) We get
$$\mathbf{E}(\tilde{\theta}) = \mathbf{E}(2/X) = \frac{2\theta^3}{2}\int_0^\infty x e^{-\theta x}\,dx = \theta^3\theta^{-2}\Gamma(2) = \theta$$
so $\tilde{\theta}$ is an unbiased estimator of θ. To get the variance we use
$$\mathbf{E}(\tilde{\theta}^2) = 2\theta^3\int_0^\infty e^{-\theta x}\,dx = 2\theta^2$$
so
$$\mathbf{V}(\tilde{\theta}) = 2\theta^2 - \theta^2 = \theta^2.$$

   (c) To find the Fisher information $i(\theta)$ for θ we use
$$\frac{\partial}{\partial\theta}\log f(x;\theta) = -x + 3/\theta, \quad -\frac{\partial^2}{\partial\theta^2}\log f(x;\theta) = 3/\theta^2 = i(\theta).$$
Thus we get
$$\mathrm{eff}(\tilde{\theta}) = \frac{\theta^2}{3\theta^2} = 1/3.$$

   (d) We get
$$\mathbf{E}(\hat{\mu}) = \frac{\theta^3}{6}\int_0^\infty x^3 e^{-\theta x}\,dx = \frac{\theta^3\Gamma(4)}{6\theta^4} = 1/\theta = \mu,$$
so $\hat{\mu}$ is an unbiased estimator of μ.

   (e) Similarly
$$\mathbf{E}(\hat{\mu}^2) = \frac{\theta^3}{18}\int_0^\infty x^4 e^{-\theta x}\,dx = \frac{\theta^3\Gamma(5)}{\theta^5} = \frac{4}{3\theta^2}$$
so
$$\mathbf{V}(\hat{\mu}) = \frac{1}{3\theta^2}.$$
To calculate the Cramér–Rao bound we find for $g(\theta) = 1/\theta$ that $g'(\theta) = -\theta^{-2}$ so the lower bound is
$$i(\theta)^{-1}\{g'(\theta)^2\} = \frac{\theta^2}{3}\theta^{-4} = \frac{1}{3\theta^2} = \mathbf{V}(\hat{\mu}).$$