TMA4295 Statistical inference
Exercise 8 - solution

Problem 1
$x \sim gamma(\alpha, \beta)$, We can notice that the gamma distribution belongs to the exponential family

$$f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta} = \frac{1}{\Gamma(\alpha)\beta^\alpha} e^{-x/\beta+(\alpha-1)\log(x)}.$$

with $c(\alpha,\beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha}$, $h(x) = 1$, $t_1(x) = \log(x)$, $w_1(\alpha,\beta) = \alpha - 1$, $t_2(x) = -x$, and $w_2(\alpha,\beta) = 1/\beta$
So we have from theorem 3.4.2

$$E(\ln(x)) = E\left(\frac{dw_1}{d\alpha} t_1(x) + \frac{dw_2}{d\alpha} t_2(x)\right) = -\frac{d\log(c(\alpha,\beta))}{d\alpha} = \frac{\Gamma(\alpha)'}{\Gamma(\alpha)} + \log(\beta).$$

Problem 2

a) By looking at the pdf of a gamma distribution we can notice that a chi-square distribution is special case of a gamma. That means that x_i are $gamma(1/2, 2)$, hence $\alpha = 1/2$, $\beta = 2$.

$$\mathrm{E}\,(x_i) = \alpha\beta = 1; \qquad \mathrm{Var}(x_i) = \alpha\beta^2 = 2$$

b) Since $x_1, x_2, ...$ is a sequence of i.i.d. variables, using the moment generating functions we can see that $z_n \sim \chi^2(n)$. Since $E(x_i) = 1$ and $\mathrm{Var}(x_i) = 2 < \infty$ we have from the central limit theorem

$$\frac{\sqrt{n}(\bar{x_n} - 1)}{\sqrt{2}} = \sqrt{n}\left(\frac{z_n}{n\sqrt{2}} - \frac{1}{\sqrt{2}}\right) \to N(0,1)$$

c) Define $g(x) = \sqrt{x} \Rightarrow g'(g) = \frac{1}{2\sqrt{x}}$.

$$g\left(\frac{1}{\sqrt{2}}\right) = \frac{1}{2^{1/4}}, \qquad g'\left(\frac{1}{\sqrt{2}}\right) = \frac{1}{2(1/\sqrt{2})^{1/2}} = \frac{1}{2^{3/2}}.$$

The delta method than gives

$$\sqrt{n}\left(W_n - \frac{1}{2^{1/4}}\right) \to N(0, 2^{-3/2}).$$

Let's observe that

$$\frac{S_n^2 n}{\sigma^2} \sim Z_n \Rightarrow S_n \sim \frac{\sigma}{\sqrt{n}}\sqrt{Z_n}$$

then

$$W_n = \frac{\sqrt{Z_n}}{\sqrt{n}2^{1/4}} \Rightarrow \sqrt{Z_n} = \sqrt{n}2^{1/4}W_n \Rightarrow S_n \sim \sigma 2^{1/4} W_n$$

$$\Rightarrow \mathrm{Var}(S_n) \approx \frac{\sigma^2 2^{1/2} 2^{-3/2}}{n} = \frac{\sigma^2}{2n}.$$

Problem 3
$X_1, ..., X_n$ i.i.d. uniformly distributed on $[0, \theta]$.

a) The moment estimator is $\hat{\theta}_M = 2\bar{X}$ and it can't be written as a function of $T(X)$.

b) If $n = 3$ the moment estimator of θ is $\hat{\theta}_M = 6$. It is not reasonable since we have an observation with value 8.

c) We first derive the MLE for θ.

$$L(\theta|\mathbf{X}) = \prod_i f(X_i|\theta) = \frac{1}{\theta^n}\prod_i I_{[0.\theta]}(X_i) = \frac{1}{\theta^n} I_{[0.\theta]}(\max_i X_i)$$

We can observe that $L(\theta|\mathbf{X})$ is a decreasing function for $\theta > \max_i X_i$, so $L(\theta|\mathbf{X})$ is maximized at $\theta = \max_i X_i$. Hence $\hat{\theta}_{MLE} = \max_i X_i$. To compute the mean, variance and MSE, we first have to find the pdf of $T = max_i X_i$. Let's first look at the cdf

$$F_T(t) = P(T \leq t) = P(X_1 \leq t, ..., X_n \leq t) = \prod_i P(X_i \leq t) = \begin{cases} 0 & t < 0 \\ \left(\frac{t}{\theta}\right)^n & 0 \leq t \leq \theta \\ 1 & t > 1 \end{cases} \quad (1)$$

and so the pdf is the derivative of 1

$$f_T(t) = \frac{nt^{n-1}}{\theta^n} \quad \text{if } 0 \leq t \leq \theta.$$

Then we easily get

$$E(T) = \frac{n}{1+n}\theta$$

$$Var(T = \frac{n}{(1+n)^2(2+n)}\theta^2$$

$$MSE(T) = Var(T) + Bias(T)^2 = \frac{2}{(1+n)(2+n)}\theta^2.$$

d) The unbiased estimator is given by $\hat{\theta} = \frac{1+n}{n}T(\mathbf{X})$, with variance $Var(\hat{\theta}) = \frac{\theta^2}{n(n+1)}$ which is also equal to the mean squares error for this estimator.

The moment estimator is unbiased with variance equal to $\frac{\theta^2}{3n}$ which is also the mean squared error.

For $n = 1$ all the estimators are the same.
For $n = 2$ $MSE(\hat{\theta}_{MLE}) = MSE(\hat{\theta}_M) > MSE(\hat{\theta})$.
For $n = 3$ $MSE(\hat{\theta}_M) > MSE(\hat{\theta}_{MLE}) > MSE(\hat{\theta})$.

Problem 4

a) We first find the pdf of $\mathbf{X}|\theta$

$$f(\mathbf{x}|\theta) = \prod_{i=1}^n f(x_i|\theta) = \frac{1}{(2\pi)^{n/2}\sigma^n}\exp\left(\frac{-\sum_i x_i^2 - n\theta^2 + 2n\theta\bar{x}}{2\sigma^2}\right).$$

Then using the fact that $f(\theta|\mathbf{X}) \propto f(\mathbf{X}|\theta)f(\theta)$ and trying to form the exponent of the form of the normal distribution we get the result.

b) The conjugate prior for the normal distribution is the normal distribution.

c) The Bayes estimator of θ is $E(\theta|\mathbf{X}) = \frac{\sigma^2}{\sigma^2+n\tau^2}m + \frac{n\tau^2}{\sigma^2+n\tau^2}\bar{\mathbf{x}}$, which is a linear combination of the prior and sample means.

From the form of the Bayes estimator it can be seen that if the prior information is unsure, i.e. τ^2 is big, then the influence of m is weak (the influence of the prior is weak). If the variance of the sample is big, i.e. σ^2 is big, the influence of the sample is weak.