This is a take home midterm. You can use your notes, books or any other source you need. You are supposed to work on your own text without external help. I'll be available to answer question in person or via email.

To solve the Exam problems, I have not collaborated with anyone and the material presented is the result of my own work.

Signature: ____________________

Name (print): ____________________

Question:	1	2	Total
Points:	65	35	100
Score:			

Question 1 .... *65 point*

Consider a random sample X_i, $i = 1, \ldots, N$, from a Normal population with expected value μ and variance σ^2, that is the X_i are i.i.d normal r.v. with $\mathbb{E}(X_i) = \mu$ and $\mathrm{var}(X_i) = \sigma^2$. In the following I'll ask you to derive confidence intervals for μ and σ under different assumption. Some of the intervals are derived in the textbook while some are not. In both case, you are asked to present a brief but complete description of the steps needed to derive the confidence interval and their meaning. In all cases, express your results in term of the inverse of the c.d.f. of the r.v. involved.

(a) (10 points) Assume that μ is unknown and σ^2 is known. Derive an exact coefficient γ confidence interval for μ using the properties of

$$\overline{X}_N = \frac{1}{N}\sum_{i=1}^{N} X_i\,.$$

This means you should find A, B such that

$$\mathbb{P}(A < \mu < B) = \gamma\,.$$

Solution:

We know that

$$Z = \sqrt{N}\frac{\overline{X}_N - \mu}{\sigma}$$

is normal standard. Thus calling $A = \overline{X}_N - a\frac{\sigma}{\sqrt{N}}$ and $B = \overline{X}_N + b\frac{\sigma}{\sqrt{N}}$ the condition becomes

$$\mathbb{P}(-a < Z < b) = \gamma\,.$$

Choosing α_1 and α_2 such that $\alpha_1 + \alpha_2 = 1 - \gamma$ we can take

$$a = -\Phi^{-1}(\alpha_1) \qquad b = -\Phi^{-1}(\alpha_2)$$

In particular if we take $\alpha_1 = \alpha_2 = (1-\gamma)/2$ we get

$$a = b = -\Phi^{-1}\left(\frac{1-\gamma}{2}\right) := z_{\frac{1-\gamma}{2}}\,.$$

Thus an exact C.I. is

$$\overline{X}_N - z_{\frac{1-\gamma}{2}}\frac{\sigma}{\sqrt{N}} \le \mu \le \overline{X}_N + z_{\frac{1-\gamma}{2}}\frac{\sigma}{\sqrt{N}}\,.$$

(b) (15 points) Assume that μ is known and σ^2 is unknown. Derive an exact coefficient γ upper confidence limit for σ^2 using the properties of

$$\Sigma_N^2 = \sum_{i=1}^{N} \left(\frac{X_i - \mu}{\sigma} \right)^2 .$$

This means you should find C such that

$$\mathbb{P}(\sigma^2 < C) = \gamma .$$

Solution: We know that Σ_N^2 has a χ^2 distribution with N d.o.f. Writing $C = c\sum_{i=1}^{N}(X_i - \mu)^2$ the condition becomes

$$\mathbb{P}\left(\Sigma_N^2 > \frac{1}{c}\right) = \gamma$$

so that we can take

$$c = \frac{1}{\chi_N^{-1}(1-\gamma)}$$

where $\chi_n(y) = \mathbb{P}(Y = y)$ is the c.d.f. of a r.v. Y with χ^2 distribution with n d.o.f. Thus an exact coefficient γ upper confidence limit is

$$\sigma^2 < \frac{1}{\chi_N^{-1}(1-\gamma)} \sum_{i=1}^{N} (X_i - \mu)^2 .$$

(c) (10 points) Assume that both μ and σ^2 are unknown. Derive an exact coefficient γ confidence interval for μ using the properties of

$$T = \frac{\sqrt{N}(\overline{X}_N - \mu)}{S_N}.$$

where

$$S_N^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(X_i - \overline{X}\right)^2$$

Solution: We know that T has a t distribution with $N-1$ d.o.f. Again we can call

$$t_{n,\alpha} = -T_n^{-1}(\alpha)$$

where $T(t)$ is the c.d.f. of a r.v. with t distribution with n d.o.f. and obtain the exact C.I.

$$\overline{X}_N - t_{N-1,\frac{1-\gamma}{2}}\frac{S_N}{\sqrt{N}} \leq \mu \leq \overline{X}_N + t_{N-1,\frac{1-\gamma}{2}}\frac{S_N}{\sqrt{N}}\,.$$

(d) (15 points) Assume again that both μ and σ^2 are unknown. Derive an exact coefficient γ upper confidence limit for σ^2 using the properties of

$$\widetilde{\Sigma}_N^2 = \sum_{i=1}^{N} \left(\frac{X_i - \overline{X}_N}{\sigma} \right)^2$$

Solution: We now have that $\widetilde{\Sigma}_N^2$ has a χ^2 distribution with $N-1$ d.o.f. so that, reasoning as before, we get the exact coefficient γ upper confidence limit

$$\sigma^2 < \frac{1}{\chi_{N-1}^{-1}(1-\gamma)} \sum_{i=1}^{N} (X_i - \overline{X}_N)^2 \,.$$

(e) (15 points) Finally, assume again that both μ and σ^2 are unknown. Derive a coefficient γ joint confidence interval for μ and upper confidence limit for σ^2 using the joint properties of

$$\overline{X}_N = \frac{1}{N}\sum_{i=1}^{N} X_i \qquad \widetilde{\Sigma}_N^2 = \sum_{i=1}^{N}\left(\frac{X_i - \overline{X}_N}{\sigma}\right)^2 .$$

This means you should find A, B and C such that

$$\mathbb{P}(A < \mu < B \ \& \ \sigma^2 < C) \geq \gamma \, .$$

Observe that there may be more than one choice for A, B and C and the interval may not be exact (see definition 8.5.1).

Solution: Since $\overline{X}_N$ and $\widetilde{\Sigma}_N^2$ are independent we can fix γ_1 and γ_2 such that $\gamma_1\gamma_2 = \gamma$ and obtain

$$\mathbb{P}\left(\overline{X}_N - z_{\frac{1-\gamma_1}{2}}\frac{\sigma}{\sqrt{N}} \leq \mu \leq \overline{X}_N + z_{\frac{1-\gamma_1}{2}}\frac{\sigma}{\sqrt{N}}\right) \cdot$$
$$\mathbb{P}\left(\sigma^2 < \frac{1}{\chi^{-1}_{N-1}(1-\gamma_2)}\sum_{i=1}^{N}(X_i - \overline{X}_N)^2\right) = \gamma$$

Since σ is unknown we can only say that

$$A = \overline{X}_N - \frac{z_{\frac{1-\gamma_1}{2}}}{\sqrt{N}}\sqrt{\sum_{i=1}^{N}\frac{(X_i - \overline{X}_N)^2}{\chi^{-1}_{N-1}(1-\gamma_2)}}$$
$$B = \overline{X}_N + \frac{z_{\frac{1-\gamma_1}{2}}}{\sqrt{N}}\sqrt{\sum_{i=1}^{N}\frac{(X_i - \overline{X}_N)^2}{\chi^{-1}_{N-1}(1-\gamma_2)}}$$
$$C = \frac{(X_i - \overline{X}_N)^2}{\chi^{-1}_{N-1}(1-\gamma_2)}$$

Question 2 . . . . *35 point*

Let X_i be a random sample of size N from a uniform distribution in $[0, A]$. Consider the statistics:

$$R = \max_{i=1,\dots,N} X_i \,.$$

(a) (10 points) Find the p.d.f. of R. Use that $R < r$ if and only if for all i, $X_i < r$.

Solution: The p.d.f. of X_i is

$$f(x) = \frac{1}{A} \quad \text{if} \quad 0 \le x \le A$$

so that

$$\mathbb{P}(R \le r) = \prod_{i=1}^{N} \mathbb{P}(X_i \le r) = \left(\frac{r}{A}\right)^N \,.$$

Finally the p.d.f. of R is

$$f_R(r) = \frac{N r^{N-1}}{A^N}$$

(b) (10 points) Find c such that $\hat{A} = cR$ is an unbiased estimator for A, that is such that $\mathbb{E}(\hat{A}) = A$ for every $A > 0$.

Solution: We have that

$$\mathbb{E}(\hat{A}) = c\mathbb{E}(R) = c\int_0^A \frac{Nr^N}{A^N} dr = c\frac{N}{N+1}A$$

so that we must take

$$c = \frac{N+1}{N}\,.$$

(c) (15 points) Find c such that the mean square error (MSE) of $\hat{A} = cR$ as an estimator of A is minimal, where

$$\mathrm{MSE}(\hat{A}) = \mathbb{E}((\hat{A} - A)^2)$$

Solution: Expanding we get

$$\mathrm{MSE}(\hat{A}) = c^2\mathbb{E}(R^2) - 2cA\mathbb{E}(R) + A^2$$

so that the minimum if for

$$c = \frac{A\mathbb{E}(R)}{\mathbb{E}(R^2)}\,.$$

Since

$$\mathbb{E}(R^2) = \int_0^A \frac{Nr^{N+1}}{A^N} dr = \frac{NA}{N+2}$$

we get

$$c = \frac{N+2}{N+1}\,.$$