MATH50011 Statistical Modelling 1
Midterm Solutions

1. (a) Provide the definition of pivotal quantity. (2 marks)
Solution: A pivotal quantity for a parameter θ is a function $t(Y,\theta)$ of the data and θ (and not any further unknown parameters) s.t. the distribution of $t(Y,\theta)$ is known.

(b) Let T_n, $n \in \mathbb{N}$, be a sequence of estimators for a parameter $\theta \in \mathbb{R}$ such that $MSE_\theta(T_n) \to 0$ as $n \to \infty$. Prove that T_n is consistent (for this question only: if you use any results from the lectures you will need to prove them). (2 marks)
Solution: since $MSE_\theta(T_n) \to 0$ as $n \to \infty$ implies that T_n is asymptotically unbiased and that $Var(T_n) \to 0$ as $n \to \infty$, then by a result of Lecture 4 we obtain that T_n is consistent. For a proof see the slides of Lecture 4.

For the remaining questions of this problem consider the following setting. Let $X_1, X_2, ...$ be a sequence of iid Exponential random variables with unknown parameter $\lambda \in (0,\infty)$. Recall that the pdf of an Exponential random variable is $f(x) = \lambda e^{-\lambda x}$ for $x \geq 0$ and $f(x) = 0$ for $x < 0$.

(c) For fixed $n \in \mathbb{N}$, compute the maximum likelihood estimator (MLE) for $\theta := \frac{1}{\lambda}$ based on the sample $X_1, ..., X_n$. Denote this estimator by $\hat{\theta}_n$. (2 marks)
Solution: From Exercise 1 of Problem sheet 3 we have that the MLE for λ based on the sample $X_1, ..., X_n$ is $\hat{\lambda}_n = \frac{1}{\bar{X}}$. Then, by functional invariance of the MLE using the bijective function $g(x) = \frac{1}{x}$ on $(0,\infty)$ we have that $\hat{\theta}_n = \bar{X}$.

(d) Show whether or not $\hat{\theta}_n$ is an unbiased and consistent estimator for θ. (1 mark)
Solution: We have that $E[\hat{\theta}_n] = E[\bar{X}] = E[X_1] = \frac{1}{\lambda} = \theta$. Hence, $\hat{\theta}_n$ is unbiased. Moreover, by the weak law of large numbers $\bar{X} \to_p E[X_1] = \theta$. Hence, $\hat{\theta}_n$ is also consistent.

(e) Let A_n, $n \in \mathbb{N}$, be a sequence of events such that $P(A_n) = \frac{1}{n}$, for every $n \in \mathbb{N}$. Let $Y_n = n\mathbf{1}_{A_n}$. Assume that $Y_1, ..., Y_n$ are independent from $X_1, ..., X_n$, for every $n \in \mathbb{N}$. Is $Y_n\hat{\theta}_n$ a consistent estimator for θ? Is it unbiased? Is it asymptotically Normal? Explain your answers in detail. (3 marks)
Solution: First, observe that $E[Y_n] = nE[\mathbf{1}_{A_n}] = nP(A_n) = 1$ for every $n \in \mathbb{N}$. Hence, by independence $E[Y_n\hat{\theta}_n] = E[Y_n]E[\hat{\theta}_n] = \theta$ and so $Y_n\hat{\theta}_n$ is an unbiased estimator for θ. Second, observe that $Y_n \to_p 0$ as $n \to \infty$ because for every $\varepsilon > 0$ we have $P(n\mathbf{1}_{A_n} > \varepsilon) = P(A_n) = \frac{1}{n}$ where the first equality comes from the fact that if $\omega \in A_n$ then $n\mathbf{1}_{A_n}(\omega) = n$ and if $\omega \notin A_n$ then $n\mathbf{1}_{A_n}(\omega) = 0$. Further, since $\hat{\theta}_n \to_p \theta$ and $Y_n \to_p 0$, by Slutsky's lemma we conclude that $Y_n\hat{\theta}_n \to_p 0$ as $n \to \infty$. Hence, $Y_n\hat{\theta}_n$ is not consistent. Now, assume that $Y_n\hat{\theta}_n$ is asymptotically Normal, then as we have seen in class and in the problem sheets, we would have that $Y_n\hat{\theta}_n$ is consistent. However, $Y_n\hat{\theta}_n$ is not consistent and so it cannot be asymptotically Normal.

2. Let $X_1, X_2, ...$ be a sequence of iid Normal random variables with known mean μ and unknown variance $\sigma^2 > 0$.

(a) Provide the definition of type 1 error and of type 2 error. (3 marks)
Solution: type 1 error is the error of rejecting H_0 when H_0 is true, while type 2 error is the error of not rejecting H_0 when H_0 is false.

(b) Show whether or not $T_n = \frac{1}{n}\sum_{i=1}^n X_i^2 - \mu^2$ is an unbiased and consistent estimator for σ^2. (2 marks)

Solution: We have that $E[\frac{1}{n}\sum_{i=1}^n X_i^2 - \mu^2] = E[X_1^2] - \mu^2 = Var(X_1) = \sigma^2$. Hence, it is unbiased. Moreover, by weak law of large numbers we have that $\frac{1}{n}\sum_{i=1}^n X_i^2 \to E[X_1^2]$ and so $\frac{1}{n}\sum_{i=1}^n X_i^2 - \mu^2 \to E[X_1^2] - \mu^2 = \sigma^2$. Thus, it is also consistent.

(c) Let $\mu = 0$ and let $\sigma_0^2 > 0$. Build an exact test of level $\alpha = 0.05$ based on T_n for $H_0 : \sigma^2 = \sigma_0^2$ vs $H_1 : \sigma^2 > \sigma_0^2$. (2 marks)

Solution: Since $\mu = 0$ we have that, under H_0, $\frac{X_1}{\sigma_0} \sim N(0,1)$ and so $\frac{X_1^2}{\sigma_0^2} \sim \chi_1^2$. Using that sum of n independent χ_1^2 is a χ_n^2 (see also Exercise 3 of Problem sheet 4), we have $nT_n/\sigma_0^2 \sim \chi_n^2$. Thus, the test of level α based on T reject H_0 if $nT_n/\sigma_0^2 > c$, where c satisfies $P(Z > c) = 0.05$ for $Z \sim \chi_n^2$.

(d) Construct the power function of the test built in point (c), making explicit the dependence on the parameter, and draw it. (2 marks)

Solution: The power function is

$$\beta(\sigma^2) = P_{\sigma^2}(nT_n/\sigma_0^2 > c) = P_{\sigma^2}(nT_n/\sigma^2 > c\sigma_0^2/\sigma^2) = P(Z > c\sigma_0^2/\sigma^2)$$

where $Z \sim \chi_n^2$. The power function is an increasing function of σ^2.

(e) Explain how your answers for points (c) and (d) would change if we had $H_0 : \sigma^2 \leq \sigma_0^2$ instead of $H_0 : \sigma^2 = \sigma_0^2$. (1 mark)

Solution: It would not change because the rejection region in both cases is the same.

(Total 20 marks)