Solutions finals STK4011-f18/STK9011-f18

Problem 1

a) (i) The probability mass function $f(x|p) = p^x(1-p)^{1-x}$ can be written $f(x|p) = \exp(x\log(p/(1-p))(1-p)$ which has the form of a class of distributins belonging to the exponential family of distributions by setting $h(x) = 1, c(p) = 1 - p, t(x) = x$ and $w(p) = \log(p/(1-p))$.

(ii) A statistic $T(X_1, \ldots, X_n)$ is sufficient with respect to the parameter p if the conditional distribution $P_p(X_1 = x_1, \ldots, X_n = x_n | T(X_1, \ldots, X_n) = t)$ does not depend on p. For the Bernoulli distributed variables

$$P_p(X_1 = x_1, \ldots, X_n = x_n | \Sigma_{i=1}^n x_i = x) = \frac{P_p(X_1=x_1,\ldots,X_n=x_n,\Sigma_{i=1}^n x_i = x)}{P_p(\Sigma_{i=1}^n x_i = x)}$$

$$= \begin{cases} \frac{p^{\Sigma x_i}(1-p)^{n-\Sigma x_i}}{\binom{n}{x}p^x(1-p)^{n-x}} & \text{if } \Sigma x_i = x \\ 0 & \text{else} \end{cases}$$

$$= \begin{cases} \frac{1}{\binom{N}{x}} & \text{if } \Sigma x_i = x \\ 0 & \text{else} \end{cases}$$

which shows that the sum $\Sigma_{i=1}^n X_i$ is sufficient for p. Alternatively one can infer the sufficiency of $\Sigma_{i=1}^n X_i$ from the general properties of distributions of belonging to the exponential family, since the Bernoulli(p) distribution does. Use of the factorization theorem is also an option.

The sum $\Sigma_{i=1}^n X_i$ is minimal sufficient for p if for sample points $(x_1 \ldots, x_n)$ and $(y_1 \ldots, y_n)$ the ratio

$$\frac{P_p(X_1 = x_1, \ldots, X_n = x_n)}{P_p(X_1 = y_1, \ldots, X_n = y_n)}$$

does not depend on p if and only if $\Sigma_{i=1}^n x_i = \Sigma_{i=1}^n y_i$. But

$$\frac{P_p(X_1 = x_1, \ldots, X_n = x_n)}{P_p(X_1 = y_1, \ldots, X_n = y_n)} = p^{\sum_{i=1}^n x_i - \sum_{i=1}^n y_i}(1-p)^{\sum_{i=1}^n y_i - \sum_{i=1}^n x_i}$$

does not depend on p if and only if $\sum_{i=1}^n x_i = \sum_{i=1}^n y_i$ so $\sum_{i=1}^n X_i$ is minimal sufficient for p.

(iii) The Bernoulli(p) distribution belongs to the exponential family of distributions. Also the parameter set $0 < p < 1$ contains an open set. Then the sufficient statistic $T = \sum_{i=1}^n X_i$ is complete. Alternatively this follows from $E_p[g(x)] = \sum_{x=1}^n g(x)\binom{n}{x}p^x(1-p)^{n-x} = (1-p)^n \sum_{x=1}^n g(x)\binom{n}{x}(\frac{p}{(1-p)})^x$ being a polynomial in $p/(1-p)$ so that $E_p[g(x)] = 0 \ \forall \ p$ implies that $g(x) = 0, \ x = 0, \ldots, n$

b) (i)The likelihood is $L(p|x_1, \ldots, x_n) = \prod_{i=1}^n p^{x_i}(1-p)^{1-x_i}$, so

$$l(p|x_1, \ldots, x_n) = \log L(p|x_1, \ldots, x_n) = (\log p)\sum_{i=1}^n x_i + (n - \sum_{i=1}^n x_i)\log(1-p).$$

The first order condition is therefore

$$\frac{\partial l(p|x_1,\ldots,x_n)}{\partial p} = \frac{\sum_{i=1}^n x_i}{p} - \frac{n-\sum_{i=1}^n x_i}{1-p} = 0$$

with solution $\hat{p} = \frac{1}{n}\sum_{i=1}^n x_i$. Also $\frac{\partial l(p|x_1,\ldots,x_n)}{\partial p}$ change sign from positive to negative at $\frac{1}{n}\sum_{i=1}^n x_i$ so it correspods to a maximum, and $\hat{p} = \frac{1}{n}\sum_{i=1}^n x_i$ is the maximum likelihood estimator.

(ii)$E[X_i] = p$ and $Var(X_i) = p(1-p)$. Therefore the the central limit theorem implies that $\sqrt{n}\frac{\frac{1}{n}\sum_{i=1}^n X_i - p}{\sqrt{p(1-p)}} = \sqrt{n}\frac{(\hat{p}-p)}{\sqrt{p(1-p)}}$ converges in distribution toward the distribution of a standard normally distributed variable.

c) (i) By the invariance property of the maximum likelihood estimator, the maximum likelihood estimator of $\tau(p)$ is $\tau(\hat{p}) = \hat{p}(1-\hat{p}) = g(\hat{p})$.

(ii)Since $g'(p) = 1-2p$, it follows using the delta method when $g'(p) \neq 0$, or $p \neq 1/2$, that

$$\sqrt{n}(g(\hat{p}) - g(p)) = \sqrt{n}g'(p)(\hat{p}-p) + R_n = \sqrt{n}(1-2p)(\hat{p}-p) + R_n$$

where R_n converges toward 0 in probability. By using Slutsky's theorem it follows that $\sqrt{n}(\hat{p}(1-\hat{p}) - p(1-p))$ converges in distribution toward the distribution of a $n(0,(1-2p)^2 p(1-p))$ distributed variable.

When $p = 1/2$, $(n[\hat{p}(1-\hat{p}) - p(1-p)]$ converges in distribution toward the distribution of a $\frac{p(1-p)}{2}g''(p)\chi_1^2 = \frac{1}{4}\frac{1}{2}(-2)\chi_1^2 = -\frac{1}{4}\chi_1^2$ distributed random variable.

d) Let $\tilde{p} = I(X_1 = 1 \text{ and } X_2 = 0)$. Then $E(\tilde{p}) = P(X_1 = 0 \text{ and } X_2 = 0) = p(1-p)$ since X_1 and X_2 are independent, so $\tilde{p}$ is unbiased. From part a) $\sum_{i=1}^n X_i$ is sufficient for p. Hence $p^* = E[\tilde{p}|\sum_{i=1}^n X_i]$ is an unbiased estimator. But

$$\begin{aligned}
E[\tilde{p}|\sum_{i=1}^n X_i = x] &= P(X_1 = 1, X_2 = 0, \sum_{i=3}^n X_i = x-2)/P(\sum_{i=1}^n X_i = x) \\
&= p(1-p)\binom{n-2}{x-1}p^{x-1}(1-p)^{n-2-(x-1)}/\binom{n}{x}p^x(1-p)^{n-x} \\
&= \binom{n-2}{x-1}p^x(1-p)^{n-x}/\binom{n}{x}p^x(1-p)^{n-x} = \binom{n-2}{x-1}/\binom{n}{x} \\
&= \frac{x(n-x)}{n(n-1)}
\end{aligned}$$

Thus $p^* = \frac{\sum_{i=1}^n X_i(n-\sum_{i=1}^n X_i)}{n(n-1)} = \hat{p}(1-\hat{p})\frac{n}{n-1}$. Since $\sum_{i=1}^n X_i$ is sufficient for p and complete, p^* is the best unbiased estimator.

e) Consider the case $p \neq 1/2$. But $\sqrt{n}(p^* - p(1-p)) = \sqrt{n}[\frac{n}{n-1}\hat{p}(1-\hat{p}) - p(1-p)] = \sqrt{n}[\hat{p}(1-\hat{p}) - p(1-p)] - \sqrt{n}\frac{1}{n-1}p(1-p)$. Since $\sqrt{n}\frac{1}{n-1}p(1-p) \to 0$, it follows by Slutsky's theorem that $\sqrt{n}(p^* - p(1-p))$ and $\sqrt{n}[\hat{p}(1-\hat{p}) - p(1-p)]$ have the same limits in distribution. Hence the asymptotic variances are the same so the asymptotic relative efficiency of the BUE of $p(1-p)$ with respect to the MLE is 1.

Problem 2

a) The simultanous probability density is

$$\begin{aligned}\prod_{i=1}^n f(x_i|\mu) &= (\frac{1}{\sqrt{2\pi}})^n \exp(-\frac{1}{2}\sum_{i=1}^n (x_i-\mu)^2) \\ &= (\frac{1}{\sqrt{2\pi}})^n \exp(-\frac{1}{2}\sum_{i=1}^n x_i^2) \exp(-\frac{n}{2}\mu^2) \exp((\sum_{i=1}^n x_i)\mu).\end{aligned}$$

The ratio

$$\frac{\prod_{i=1}^n f(x_i|\mu)}{\prod_{i=1}^n f(y_i|\mu)} = \frac{\exp(-\frac{1}{2}\sum_{i=1}^n x_i^2)}{\exp(-\frac{1}{2}\sum_{i=1}^n y_i^2)} \exp((\sum_{i=1}^n x_i - \sum_{i=1}^n y_i)\mu)$$

is constant in μ if and only if $\sum_{i=1}^n x_i = \sum_{i=1}^n y_i$ so $\sum_{i=1}^n X_i$ is a minimal sufficient statistic.

b) The logarithm of the likelihood is $\log L = -\frac{n}{2}\log 2\pi - \frac{1}{2}\sum_{i=1}^n (x_i-\mu)^2$, $\frac{\partial \log L}{\partial \mu} = \sum_{i=1}^n (x_i - \mu)$ and $\frac{\partial^2 \log L}{\partial \mu^2} = -n$. Hence the Cramer-Rao lower bound is

$$\frac{(\partial \mu^2)^2}{-\frac{\partial^2 \log L}{\partial \mu^2}} = \frac{4\mu^2}{n}.$$

c) $E[\bar{X}^2] = Var(\bar{X}) + (E[\bar{X})^2 = \frac{1}{n} + \mu^2$, so $\bar{X}^2 - \frac{1}{n}$ is an unbiased estimator for μ^2, and $\bar{X} = \sum_{i=1}^n X_i/n$ is sufficient for μ. Completeness follows since the distribution of $\bar{X}$ belongs to the exponential family of distributions where the parameter set contains an open set. Hence $\bar{X}^2 - \frac{1}{n}$ is an best unbiased or UMVUE estimator.

Problem 3

a) The simultanous probability density is

$$\prod_{i=1}^n f(x_i|\mu) = (\frac{1}{\sqrt{2\pi}\sigma})^n \exp(-\frac{1}{2\sigma^2}\sum_{i=1}^n x_i^2).$$

From the Neyman-Pearson lemma it follows that the test function for the UMP test must satisfy

$$\phi = \begin{cases} 1 & \text{if} \quad \frac{(\frac{1}{\sqrt{2\pi}\sigma_1})^n \exp(-\frac{1}{2\sigma_1^2}\sum_{i=1}^n x_i^2)}{(\frac{1}{\sqrt{2\pi}\sigma_0})^n \exp(-\frac{1}{2\sigma_0^2}\sum_{i=1}^n x_i^2)} = \exp(\frac{1}{2}(\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2})\sum_{i=1}^n x_i^2) > k \\ 0 & \text{if} \quad \frac{(\frac{1}{\sqrt{2\pi}\sigma_1})^n \exp(-\frac{1}{2\sigma_1^2}\sum_{i=1}^n x_i^2)}{(\frac{1}{\sqrt{2\pi}\sigma_0})^n \exp(-\frac{1}{2\sigma_0^2}\sum_{i=1}^n x_i^2)} = \exp(\frac{1}{2}(\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2})\sum_{i=1}^n x_i^2) < k \end{cases}$$

which since $\sigma_1 > \sigma_0$ has the form

$$\phi = \begin{cases} 1 & \text{if} \quad \sum_{i=1}^n x_i^2 > c \\ 0 & \text{if} \quad \sum_{i=1}^n x_i^2 < c \end{cases}$$

for a suitable constant $c = c(k)$.

b) Since $\sum_{i=1}^n X_i^2/\sigma^2$ is χ_n^2 distributed the test has size α if $\alpha = P_{\sigma_0}(\sum_{i=1}^n X_i^2 > c) = P_{\sigma_0}(\sum_{i=1}^n X_i^2/\sigma_0^2 > c/\sigma_0^2)$. Therefore if $\chi_{n,\alpha}$ is the $1-\alpha$ quantile in the χ_n^2-distribution, $c/\sigma_0^2 = \chi_{n,\alpha}$ or $c = \sigma_0^2\chi_{n,\alpha}$.

c) One possibility is to use the Karlin-Rubin theorem which says that since $\sum_{i=1}^n X_i^2/\sigma^2$ is sufficient, which follows from the factorization theorem, and has monotone likelihood ratio, the test with rejection region $\sum_{i=1}^n X_i^2 > c$ is UMP at its level for testing $H_0' : \sigma = \sigma_0$ versus the alternative $H_1' : \sigma = \sigma_1$ where $\sigma_1 > \sigma_0$.

   A more direct argument runs as follow noticing that the test does not depend on σ_1^2. The probability of rejection is $P(\sum_{i=1}^n X_i^2 > \sigma_0^2\chi_{n,\alpha}) = P(\sum_{i=1}^n X_i^2/\sigma^2 > \sigma_0^2\chi_{n,\alpha}/\sigma^2) = 1 - F(\sigma_0^2\chi_{n,\alpha}/\sigma^2)$ where F is the cumulative distribution function of a χ_n^2 variable. $1-F(\sigma_0^2\chi_{n,\alpha}/\sigma^2)$ is increasing in σ^2. Hence the test from part a) has level α also for the hypothesis test $H_0 : \sigma \le \sigma_0$ versus $H_1 : \sigma > \sigma_0$. Such tests are a subset of tests having level α for $H_0 : \sigma = \sigma_0$. Since the test is UMP in this set and belongs to the subset it must also be UMP in the subset. Hence the test is UMP for $H_0 : \sigma \le \sigma_0$ versus $H_1 : \sigma > \sigma_0$.

d) (i) The power function is the probability of rejection which was calculated in part c) and is $\beta(\sigma^2) = 1-F(\sigma_0^2\chi_{n,\alpha}/\sigma^2)$ where F is the cumulative distribution function of a χ_n^2 variable.

   (ii) It is increasing in σ^2 and $\beta(\sigma_0^2) = \alpha$, so $\beta(\sigma^2) > \alpha$ when $\sigma^2 > \sigma_0^2$ which is the definition of an unbiased test.

Problem 4

a) (i)An ancillary statistics is a statistic S(X) whose distribution does not depend on the parameter θ. (ii)Basu's theorem states that if T(X) is a complete and minimal sufficient statistic, then T(X) is independent of any ancillary statistic.

b) $X_1 - \bar{X} = (1 - \frac{1}{n})X_1 - \frac{1}{n}\sum_{i=2}^{n} X_i$. Hence $E(X_1 - \bar{X}) = \mu - \mu = 0$ and $Var(X_1 - \bar{X}) = (1 - \frac{1}{n})^2 + \frac{(n-1)}{n^2} = \frac{(n-1)^2 + (n-1)}{n^2} = \frac{(n-1)n}{n^2} = \frac{n-1}{n}$ since $X_1, \ldots, X_n$ are i.i.d. $n(0,1)$ variables and linear combinations of normally distributed random variables are normally distributed. Hence $X_1 - \bar{X} \sim n(0, \frac{n-1}{n})$ which do not depend on μ, and therefore $X_1 - \bar{X}$ is ancillary.

c) That $\bar{X}$ is minimal sufficient and that the distributions belong to the exponential family was argued in problem 1, part a). Also the set of parameter contains an open set, so $\bar{X}$ is complete. That $X_1 - \bar{X}$ is ancillary was explained in part b). Hence the assumptions of Basu's theorem is satisfied so $\bar{X}$ and $X_1 - \bar{X}$ must be independent.