Comments and corrections to r.samworth@statslab.cam.ac.uk

1. Let $X \sim N_d(\mu, \Sigma)$, and let A be an arbitrary $p \times d$ matrix. Give two proofs of the fact that $AX \sim N_p(A\mu, A\Sigma A^T)$: one directly from the (general) definition and the other using moment generating functions.

2. If $Y \sim \chi_m^2$ and $Z \sim \chi_n^2$ are independent, we say $X = \frac{Y/m}{Z/n}$ has an F-distribution with m and n degrees of freedom, and write $X \sim F_{m,n}$. Show that the probability density function of X is

$$f_X(x; m, n) = \frac{\Gamma(\frac{m+n}{2})(m/n)^{m/2} x^{\frac{m}{2}-1}}{\Gamma(m/2)\Gamma(n/2)(1 + \frac{mx}{n})^{(m+n)/2}}, \quad x > 0.$$

3. (Law of small numbers) For each $n \in \mathbb{N}$, let $X_{n1}, \ldots, X_{nn} \overset{iid}{\sim} \text{Bernoulli}(p_n)$ and let $S_n = \sum_{i=1}^{n} X_{ni}$. Prove that if $np_n \to \lambda \in (0, \infty)$ as $n \to \infty$, then for each $x \in \{0, 1, \ldots\}$,

$$\mathbb{P}(S_n = x) \to \mathbb{P}(Y = x)$$

as $n \to \infty$, where $Y \sim \text{Poi}(\lambda)$.

4. Suppose that (X, T) have a joint probability density function $f_{X,T}(x, t; \theta)$. Prove the factorisation criterion for the sufficiency of T.

5. (a) Let $X_1, \ldots, X_n$ be independent Poisson random variables, with X_i having parameter $i\theta$ for some $\theta > 0$. Find a real-valued sufficient statistic T, and compute its distribution. Show that the maximum likelihood estimator $\hat{\theta}$ of θ is a function of T, and show that it is unbiased.

(b) For some $n > 2$, let $X_1, \ldots, X_n \overset{iid}{\sim} \text{Exp}(\theta)$. Find a real-valued sufficient statistic T, and compute its distribution. Show that the maximum likelihood estimator $\hat{\theta}$ of θ is a function of T, and is biased but asymptotically unbiased. Find a reparametrisation of the model in terms of a parameter $\psi = h(\theta)$, where h is bijective, such that the maximum likelihood estimator $\hat{\psi}$ of ψ is unbiased.

6. For some $n \geq 2$ let $X_1, \ldots, X_n \overset{iid}{\sim} U[\theta, 2\theta]$, for some $\theta > 0$. Show that $\tilde{\theta} = \frac{2}{3} X_1$ is an unbiased estimator of θ. Find an unbiased estimator $\hat{\theta}$ which is a function of a sufficient statistic taking values in a subset of $\mathbb{R}^2$ and which satisfies $\text{Var}_\theta(\hat{\theta}) < \text{Var}_\theta(\tilde{\theta})$ for all $\theta > 0$.

7. Let $X_1, \ldots, X_n \overset{iid}{\sim} U[0, \theta]$. For $\alpha \in (0, 1)$, find a $100(1-\alpha)\%$ confidence interval for θ based on a real-valued sufficient statistic.

8. Suppose that $X_1 \sim N(\theta_1, 1)$ and $X_2 \sim N(\theta_2, 1)$ independently, where θ_1 and θ_2 are unknown. Show that both the square S and circle C in $\mathbb{R}^2$, given by

$$S = \{(\theta_1, \theta_2) : |\theta_1 - X_1| \le 2.236, |\theta_2 - X_2| \le 2.236\}$$
$$C = \{(\theta_1, \theta_2) : (\theta_1 - X_1)^2 + (\theta_2 - X_2)^2 \le 5.991\}$$

are 95% confidence sets for (θ_1, θ_2). *Hint:* $\Phi(2.236) = (1 + \sqrt{0.95})/2$, *where* Φ *is the distribution function of the* $N(0, 1)$ *distribution.* What might be a sensible criterion for choosing between S and C?

9. Suppose that the number of defects on a roll of magnetic recording tape can be modelled with Poisson distribution for which the parameter λ is known to be either 1 or 1.5. Suppose the prior mass function for λ is

$$\pi_\lambda(1) = 0.4, \quad \pi_\lambda(1.5) = 0.6.$$

A random sample of five rolls of tape finds $x = (3, 1, 4, 6, 2)$ defects respectively. Show that the posterior distribution for λ given x is

$$\pi_{\lambda|X}(1|x) = 0.012, \quad \pi_{\lambda|X}(1.5|x) = 0.988.$$

10. Let $X_1, \ldots, X_n$ be independent and identically distributed with conditional probability density function $f(x|\theta) = \theta x^{\theta-1}$ $_{\{x \in (0,1)\}}$ for some $\theta > 0$. Suppose the prior distribution for θ is $\Gamma(\alpha, \lambda)$. Find the posterior distribution of θ given $X = (X_1, \ldots, X_n)$ and the Bayesian point estimator of θ under the quadratic loss function.

11. For some $n \ge 3$, let $\epsilon_1, \ldots, \epsilon_n \overset{iid}{\sim} N(0, 1)$, set $X_1 = \epsilon_1$ and $X_i = \theta X_{i-1} + (1 - \theta^2)^{1/2} \epsilon_i$ for $i = 2, \ldots, n$ and some $\theta \in (-1, 1)$. Find a sufficient statistic for θ that takes values in a subset of $\mathbb{R}^3$.

+12. Let $\hat{\theta}$ be an unbiased estimator of $\theta \in \Theta \subseteq \mathbb{R}$ satisfying $\mathbb{E}_\theta(\hat{\theta}^2) < \infty$ for all $\theta \in \Theta$. We say $\hat{\theta}$ is a *uniform minimum variance unbiased* (UMVU) estimator if $\mathrm{Var}_\theta\, \hat{\theta} \le \mathrm{Var}_\theta\, \tilde{\theta}$ for all $\theta \in \Theta$ and any other unbiased estimator $\tilde{\theta}$. Prove that a necessary and sufficient condition for $\hat{\theta}$ to be a UMVU estimator is that $\mathbb{E}_\theta(\hat{\theta}U) = 0$ for all $\theta \in \Theta$ and all U with $\mathbb{E}_\theta(U) = 0$ and $\mathbb{E}_\theta(U^2) < \infty$ (i.e. '$\hat{\theta}$ is uncorrelated with every unbiased estimator of 0'). Is the estimator $\hat{\theta}$ in **6.** a UMVU estimator?