Solutions of the mid-term exam problems
November 11, 2004

1. (a) $\begin{pmatrix} X_1 \\ Z_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} X_1 \\ Y_1 \end{pmatrix}$ so that $\begin{pmatrix} X_1 \\ Z_1 \end{pmatrix} \sim N\left(\begin{pmatrix} 1 \\ 5 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 2 & 6 \end{pmatrix}\right)$ and $\mathrm{Cov}(X_1, Z_1) = 2$ which implies that X_1, Z_1 are not independent. In the same way: $\begin{pmatrix} Y_1 \\ Z_1 \end{pmatrix} \sim N\left(\begin{pmatrix} 2 \\ 5 \end{pmatrix}, \begin{pmatrix} 3 & 0 \\ 0 & 6 \end{pmatrix}\right)$, thus $\mathrm{Cov}(Y_1, Z_1) = 0$ which implies that Y_1, Z_1 are independent.
$\begin{pmatrix} Z_1 \\ V_1 \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ 2 & -3 \end{pmatrix} \begin{pmatrix} X_1 \\ Y_1 \end{pmatrix}$, thus $\begin{pmatrix} Z_1 \\ V_1 \end{pmatrix} \sim N\left(\begin{pmatrix} 5 \\ -4 \end{pmatrix}, \begin{pmatrix} 6 & 4 \\ 4 & 43 \end{pmatrix}\right)$.
The vector $(X_1, Z_1, V_1)^T$ is normal since it is a linear transformation of a normal vector but it does not have a density because its coordinates are linearly dependent: $V_1 = -3Z_1 + 11X_1$. The covarince matrix of this vector must be singular (one can also show this by direct calculations) and therefore the density does not exist.
By the Fisher theorem, the joint distribution of $(\bar{V}_7, 6S_v^2/\mathrm{Var}(V_1))^T$ is the product of $N(-4, \frac{43}{7})$ and χ_6^2 (since the coordinates are independent).

   (b) From (a), we know that $Z_1 \sim N(5, 6)$, $V_1 \sim N(-4, 43)$. Therefore, by the theorem of Fisher (recall also that if $Y \sim \chi_k^2$, then $EY = k$ and $\mathrm{Var}(Y) = 2k$), it is easy to compute $\frac{1}{5}E\bar{Z}_7 = 1$, $\frac{7}{3}\mathrm{Var}(\bar{Z}_7) = 2$, $\frac{1}{2}ES_z^2 = 3$, $\mathrm{Var}(\frac{S_z^2}{\sqrt{3}}) = 4$, $\frac{5}{43}ES_v^2 = 5$, $\mathrm{Var}(\frac{\sqrt{18}}{43}S_v^2) = 6$.

   (c) Recall $X_1 \sim N(1, 1)$, $Y_1 \sim N(2, 3)$. Calculate further $\frac{21}{101}\mathrm{Var}(7\bar{Z}_7 - 3\bar{X}_9) = \frac{21}{101}\left(\frac{49\mathrm{Var}(Z_1)}{7} + \frac{9\mathrm{Var}(X_1)}{9} - \frac{21\cdot 2\cdot 7\mathrm{Cov}(Z_1, X_1)}{7\cdot 9}\right) = 7$, $1 + \frac{1}{3}\mathrm{Var}(2S_y^2 - S_z^2) = 1 + \frac{1}{3}\big(4\mathrm{Var}(S_y^2) + \mathrm{Var}(S_z^2)\big) = 8$ because Y_1, Z_1 are independent, $1.8\mathrm{Cov}(X_1, V_1) = \frac{9}{5}\big(2\mathrm{Var}(X_1) - 3\mathrm{Cov}(X_1, Y_1)\big) = 9$, $10 + \mathrm{Cov}(X_1, V_5) = 10$ because X_1, V_5 are independent and $55P\big(\bar{Z}_7 > S_z\frac{0.906}{\sqrt{7}} + 5\big) = 55P\big(\frac{\sqrt{7}(\bar{Z}_7 - 5)}{S_z} > 0.906\big) = 55P\big(T > 0.906\big) = 55(1 - 0.8) = 11$.

2. (a) $ET_1 = \frac{nE\bar{X}_n + mE\bar{Y}_m}{n+m} = \mu$ en $ET_2 = \frac{\alpha nE\bar{X}_n + mE\bar{Y}}{m + \alpha n} = \mu$. Both T_1 and T_2 are unbiased. Therefore, $\mathrm{MSE}(T_1) = \mathrm{Var}(T_1) = \frac{n^2\mathrm{Var}(\bar{X}) + m^2\mathrm{Var}(\bar{Y})}{(n+m)^2} = \frac{(n+\alpha m)\sigma^2}{(n+m)^2}$, $\mathrm{MSE}(T_2) = \mathrm{Var}(T_2) = \frac{\alpha^2 n^2\mathrm{Var}(\bar{X}_n) + m^2\mathrm{Var}(\bar{Y}_m)}{(m+\alpha n)^2} = \frac{\alpha\sigma^2}{m+\alpha n}$. Comparing $\mathrm{MSE}(T_1)$ with $\mathrm{MSE}(T_2)$ boils down to comparing $(n + \alpha m)(m + \alpha n)$ with $\alpha(n + m)^2$ or $1 + \alpha^2$ with 2α. But $1 + \alpha^2 \geq 2\alpha$ (since $(1 - \alpha)^2 \geq 0$), thus $\mathrm{MSE}(T_1) \geq \mathrm{MSE}(T_2)$, which means that T_2 is more preferable.

   (b) By the CLT $Z_n = \sqrt{n}(\bar{X}_n - \mu) \xrightarrow{d} Z \sim N(0, \sigma^2)$ as $n \to \infty$ and $V_n = \sqrt{m}(\bar{Y}_m - \mu) \xrightarrow{d} V \sim N(0, \alpha\sigma^2)$ as $m \to \infty$. Since Z_n, V_n are independent, $(Z_n, V_n)^T \xrightarrow{d} (Z, V)^T$ as $n \to \infty$, where Z, V are independent and have the above marginal normal distributions (this follows immediately from the definition of the weak convergence). Now, apply the continuous mapping theorem to conclude that $\sqrt{n}(T_1 - \mu) = \frac{\sqrt{n}(\bar{X}_n - \mu)}{1 + m_n/n} + \frac{\sqrt{m_n}(\bar{Y}_m - \mu)\sqrt{n/m_n}}{1 + n/m_n} = \frac{Z_n}{1 + m_n/n} + \frac{V_m\sqrt{n/m_n}}{1 + n/m_n} \xrightarrow{d} \frac{1}{3}Z + \frac{\sqrt{2}}{3}V \sim N\big(0, \frac{(1+2\alpha)\sigma^2}{9}\big)$ and similarly $\sqrt{n}(T_2 - \mu) = \frac{Z_n\alpha}{\alpha + m_n/n} + \frac{V_m\sqrt{n/m_n}}{1 + \alpha n/m_n} \xrightarrow{d} \frac{\alpha}{\alpha+2}Z + \frac{\sqrt{2}}{\alpha+2}V = W \sim N\big(0, \frac{\alpha\sigma^2}{\alpha+2}\big)$ as $n \to \infty$. Notice in passing that $\frac{\alpha\sigma^2}{\alpha+2} \leq \frac{(1+2\alpha)\sigma^2}{9}$ for all $\alpha \in \mathbb{R}$. Apply the delta-method: $\sqrt{n}\big(\sin(T_2) - \sin(\mu)\big) \xrightarrow{d} \cos(\mu)W \sim N\big(0, \frac{\alpha\sigma^2(\cos(\mu))^2}{\alpha+2}\big)$ as $n \to \infty$.

   (c) Denote $\theta = (\mu, \sigma^2)$. The loglikelihood function is $l_{\mu,\sigma^2}(X, Y) = l_\theta(X, Y) = \log p_\theta(X, Y) = -\frac{(n+m)\log(2\pi) + m\log\alpha}{2} - \frac{(n+m)\log\sigma^2}{2} - \sum_{i=1}^n \frac{(X_i - \mu)^2}{2\sigma^2} - \sum_{j=1}^m \frac{(Y_j - \mu)^2}{2\alpha\sigma^2}$. Solve the likelihood equations $\frac{\partial l_\theta(X,Y)}{\partial\theta} = 0$ and derive the MLE $\hat{\mu} = T_2$ and $\hat{\sigma^2} = \frac{\alpha\sum_{i=1}^n (X_i - T_2)^2 + \sum_{j=1}^m (Y_j - T_2)^2}{\alpha(n+m)}$. The maximum of $l_\theta(X, Y)$ is achieved in this point: for any fixed σ^2 the maximum over μ

is in $\hat{\mu}$ and the maximum of $l_{\hat{\mu},\sigma^2}(X,Y)$ is in $\hat{\sigma^2}$.

If σ^2 is known, the Fisher information about μ is $I_\mu = -E_\theta\Big(\frac{\partial^2 l_{\mu,\sigma^2}(X,Y)}{\partial\mu^2}\Big) = \frac{\alpha n+m}{\alpha\sigma^2}$. The MLE T_2 is unbiased and $\mathrm{Var}(T_2) = \frac{\alpha\sigma^2}{m+\alpha n} = \frac{1}{I_\mu}$, that is, the Cramér-Rao bound is sharp.

If μ is known, the Fisher information about σ^2 is $I_{\sigma^2} = -E_\theta\Big(\frac{\partial^2 l_{\mu,\sigma^2}(X,Y)}{\partial(\sigma^2)^2}\Big) = \frac{n+m}{2\sigma^4}$. The MLE of σ^2 is $\tilde{\sigma^2} = \frac{\alpha\sum_{i=1}^n (X_i-\mu)^2+\sum_{j=1}^m (Y_j-\mu)^2}{\alpha(n+m)}$ which is unbiased $E\tilde{\sigma^2} = \sigma^2$ and $\mathrm{Var}(\tilde{\sigma^2}) = \frac{\alpha^2\sum_{i=1}^n \mathrm{Var}(X_i-\mu)^2+\sum_{j=1}^m \mathrm{Var}(Y_j-\mu)^2}{\alpha^2(n+m)^2} = \frac{2\sigma^4}{n+m} = \frac{1}{I_{\sigma^2}}$, that is, the Cramér-Rao bound is sharp.

(d) Let $T_3 = a\bar{X} + b\bar{Y}$ be an unbiased estimator of μ. We must have $ET_3 = \mu$ or $a\mu + b\mu = \mu$, or $a + b = 1$. Thus $a = 1 - b$ and $\mathrm{MSE}(T_3) = \mathrm{Var}(T_3) = (a^2 + \alpha b^2)\sigma^2/n$. We have to show that $\mathrm{MSE}(T_2) \le \mathrm{MSE}(T_3)$ or $\frac{\alpha}{1+\alpha} \le (1-b)^2 + \alpha b^2 = 1 - 2b + (1+\alpha)b^2$, or equivalently $0 \le 1-2b(1+\alpha)+(1+\alpha)^2b^2 = (1-(1+\alpha)b)^2$, which is always true. The proof is completed.

3. (a) The moment estimator is a solution of the equations $\bar{X}_n = EX_1 = E(Y+\theta_2) = \theta_1+\theta_2$, $\frac{1}{n}\sum_{i=1}^n X_i^2 = \overline{X^2}_n = EX_1^2 = E(Y+\theta_2)^2 = (\theta_1+\theta_2)^2 + \theta_1^2$. This yields $\tilde{\theta}_1 = \sqrt{\overline{X^2}_n - \bar{X}_n^2}$ and $\tilde{\theta}_2 = \bar{X}_n - \sqrt{\overline{X^2}_n - \bar{X}_n^2}$. Notice that $\tilde{\theta}_1 + \tilde{\theta}_2 = \bar{X}_n$ and $\mathrm{Var}(X_1) = \mathrm{Var}(Y) = \theta_1^2$. So, by the CLT, $\sqrt{n}\big(\tilde{\theta}_1+\tilde{\theta}_2-(\theta_1+\theta_2)\big) = \sqrt{n}(\bar{X}_n - EX_1) = Y_n \xrightarrow{d} Y \sim N(0,\theta_1^2)$ as $n\to\infty$. By the delta-method, with $\phi(x) = x^{-1}$, $\sqrt{n}\big((\tilde{\theta}_1+\tilde{\theta}_2)^{-1} - (\theta_1+\theta_2)^{-1}\big) = \sqrt{n}\big(\phi(\bar{X}_n) - \phi((\theta_1+\theta_2)\big) \xrightarrow{d} \phi'(\theta_1+\theta_2)Y \sim N\big(0, \frac{\theta_1^2}{(\theta_1+\theta_2)^4}\big)$ as $n\to\infty$.

(b) Let $X_{(1)} = \min(X_1,\ldots,X_n)$, $\theta = (\theta_1,\theta_2)$. Write down the likelihood function $p_\theta(X) = \theta_1^{-1}\exp\big\{\frac{n\theta_2 - \sum_{i=1}^n X_i}{\theta_1}\big\} I\{\theta_2 \le X_{(1)}\}$. For any fixed $\theta_1 > 0$, it is maximized at $\hat{\theta}_2 = X_{(1)}$. Therefore $\max_{\theta_2\in\mathbb{R},\theta_1>0} p_\theta(X) = \max_{\theta_1>0} p_{\hat{\theta}_2,\theta_1}(X)$. The maximum of $p_{\hat{\theta}_2,\theta_1}(X)$ is attained at a solution of the equation $\frac{\partial \log p_{\hat{\theta}_2,\theta_1}(X)}{\partial\theta_1} = 0$ or $-\frac{n}{\theta_1} - \frac{n\hat{\theta}_2 - \sum_{i=1}^n X_i}{\theta_1^2} = 0$, which is $\hat{\theta}_1 = \bar{X}_n - X_{(1)}$ (this gives the unique maximum). Thus, the MLE is $(\hat{\theta}_1,\hat{\theta}_2)$. The MLE is biased: $E\hat{\theta}_2 = EX_{(1)} = \theta_2 + \frac{EZ_n}{n} = \theta_2 + \frac{\theta_1}{n} \ne \theta_2$ since $Z_n \sim \mathrm{Exp}(1/\theta_1)$ by (c). The MLE is asymptotically unbiased: $E\hat{\theta}_2 = \theta_2 + \frac{\theta_1}{n} \to \theta_2$ and $E\hat{\theta}_1 = E\bar{X}_n - EX_{(1)} = \theta_1+\theta_2-(\theta_2+\frac{\theta_1}{n}) \to \theta_1$ as $n\to\infty$.

(c) Compute $F_{X_1}(x) = \int_{\theta_2}^x f_{\theta_1,\theta_2}(u)du = 1 - e^{-\frac{x-\theta_2}{\theta_1}}$ for $x \ge \theta_2$ and $F_{X_1}(x) = 0$ for $x < \theta_2$. Denote $Z_n = n(\hat{\theta}_2 - \theta_2)$, then for $x \ge 0$ $F_{Z_n}(x) = P(n(\hat{\theta}_2-\theta_2) \le x) = P\big((X_{(1)} \le \theta_2 + \frac{x}{n}\big) = 1 - \big(1 - F_{X_1}(\theta_2 + \frac{x}{n})\big)^n = 1 - e^{-\frac{x}{\theta_1}}$ and $F_{Z_n}(x) = 0$ for $x < 0$, i.e. $Z_n \sim \mathrm{Exp}(1/\theta_1)$ for all $n \in \mathbb{N}$. Thus $n(\hat{\theta}_2 - \theta_2) = Z_n \xrightarrow{d} Z \sim \mathrm{Exp}(1/\theta_1)$ as $n\to\infty$. Using this, (a) and Slutski's theorem, we obtain that $V_n = \sqrt{n}(\hat{\theta}_1 - \theta_1) = \sqrt{n}\big(\bar{X}_n - (\theta_1+\theta_2) - (X_{(1)} - \theta_2)\big) = \sqrt{n}\big(\bar{X}_n - (\theta_1+\theta_2)\big) + \frac{1}{\sqrt{n}} n(X_{(1)} - \theta_2) = Y_n + o_p(1)Z_n \xrightarrow{d} Y \sim N(0,\theta_1^2)$ as $n\to\infty$. By the continuous mapping theorem, $\sin(n(\hat{\theta}_2-\theta_2)) = \sin(Z_n) \xrightarrow{d} \sin(Z)$, with $Z \sim \mathrm{Exp}(1/\theta_1)$, and $\cos\big(n^{1/3}(\hat{\theta}_1-\theta_1)\big) = \cos\big(n^{-1/6}\sqrt{n}(\hat{\theta}_1-\theta_1)\big) = \cos(n^{-1/6}V_n) \xrightarrow{d} \cos(0) = 1$ as $n\to\infty$. Therefore, by Slutski's theorem $\frac{\sin(n(\hat{\theta}_2-\theta_2))}{\cos(n^{1/3}(\hat{\theta}_1-\theta_1))} \xrightarrow{d} \sin(Z)$ as $n\to\infty$, where $Z \sim \mathrm{Exp}(1/\theta_1)$.

(d) If θ_2 is known, the Fisher information about θ_1 is $I_{\theta_1} = \mathrm{Var}\Big(\frac{\partial\log p_{\theta_2,\theta_1}(X)}{\partial\theta_1}\Big) = \mathrm{Var}\Big(\frac{\sum_{i=1}^n X_i}{\theta_1^2}\Big) = \frac{n}{\theta_1^2}$. The MLE for θ_1 is then $\check{\theta}_1 = \bar{X}_n - \theta_2$ (this follows from (b)) which is unbiased for θ_1 and $\mathrm{Var}(\check{\theta}_1) = \frac{\mathrm{Var}(X_1)}{n} = \frac{\theta_1^2}{n} = \frac{1}{I_{\theta_1}}$, i.e. the Cramér-Rao bound is sharp.