Stat 543 Spring 2005
Assignment 6 Solutions (Mar. 3rd, 2005)

1. Suppose that I am interested in a continuous distribution with *pdf* $f(x) = K\phi(x)sin^2(x)$ for $\phi(x)$ the standard normal *pdf* and K some appropriate normalizing constant (that to my knowledge can not be evaluated "in closed form" but that MathCad indicates is no more than 2.32).

   (a) Carefully and completely describe a rejection sampling algorithm for generating X from this distribution.

   (b) Then consider the matter of finding the mean and variance for this distribution, and $P[.3 < X < 1.2]$. Describe how you would use 10,000 values simulated as in a) to evaluate these.

   Solution:

   (a) As $K \leq 2.32$, take $M = 2.32$ then $M\phi(x) > f(x) = K\phi(x)\sin^2 x$
   The rejection algorithm
   (1) Generate $X^{**} \sim N(0,1)$
   (2) Generate $U \sim N(0,1)$
   (3) If $MU\phi(x^{**}) < K\phi(x^{**})\sin^2 x^{**}$, set $x^* = x^{**}$
   otherwise, return to step (1).
   This algorithm generates $x_1^*, x_2^*, ...$

   (b) By LLN, as $n \to \infty$, $\hat{\mu} = \bar{x}_n \xrightarrow{P} \mu$, $\hat{S}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 \xrightarrow{P} \sigma^2$
   and $\widehat{P[0.3 < x < 1.2]} = \frac{1}{n}\sum_{i=1}^{n} I(0.3 < x < 1.2) \xrightarrow{P} P[0.3 < x < 1.2]$
   We will use $\hat{\mu} = \bar{x}_{1000}$
   $\widehat{s^2} = \frac{1}{(n-1)}\sum_{i=1}^{1000}(x_i - \overline{x})^2$
   and $\widehat{P}[0.3 < x < 1.2] = \frac{1}{1000}\sum_{i=1}^{1000} I(0.3 < x < 1.2)$ to evaluate μ, σ^2 and $P[0.3 < x < 1.2]$.

2. A 10-dimensional random vector X has a complicated (joint) pdf that is proportional to the function.

$$h(x) = \begin{cases} \exp\left(\prod_{i=1}^{10} \sin^2(x_i)\right) & \text{if each } x_i \in (0,1) \\ 0 & \text{otherwise} \end{cases}$$

   Notice that $h(x) \leq e^1$.

   (a) Carefully describe an algorithm (using only iid Uniform (0, 1) random for generating a (10-dimensional) realization of M.

   (b) Describe a simulation-based method of approximating EX_1X_2.

   Solution:

(a) As $h(\underline{x}) \le e^1 \cdot 1$, take $M = e^1$, then $M \cdot 1 \ge \exp\left(\prod_{i=1}^{10} \sin^2 x_i\right)$

The rejection algorithm

(1) Generate $X_j^{**} \sim U(0,1), j = 1, ..., 10$

(2) Generate $U \sim U(0,1)$.

(3) If $MU < \exp\left(\prod_{j=1}^{10} \sin^2 x_j^{**}\right)$ set $\underline{x}_i^* = \underline{x}_i^{**}, \underline{x}_i^* = (x_{i1}^{**}, ..., x_{i10}^{**})$

otherwise, return to step (1).

This algorithm generates $x_1^*, x_2^*, ...$

(b) As $\frac{\sum_{i=1}^n x_{i_1}^* x_{i_2}^*}{n} \xrightarrow{P} EX_1X_2$, By *LLN*.

We use $\hat{\mu}_{12} = \frac{1}{n}\sum_{i=1}^n x_{i_1}^* x_{i_2}^*$ to evaluate EX_1X_2.

3. Consider the following model. Given parameters $\lambda_1, ..., \lambda_N$ variables $X_1, ..., .X_N$ are independent Poisson variables, $X_i \sim \text{Poisson}(\lambda_i)$. M is a parameter taking values in $\{1, 2, ..., N\}$ and if $i \le M, \lambda_i = \mu_1$, while if $i > M, \lambda_i = \mu_2$. (M is the number of Poisson means that are the same as that of the first observation.) With parameter vector $\theta = (M, \mu_1, \mu_2)$ belonging to $\Theta = \{1, 2, ..., N\} \times (0, \infty) \times (0, \infty)$ we wish to make inference about M based on $X_1, ..., .X_N$ in a Bayesian framework. As matters of notational convenience, let $S_m = \sum_{i=1}^m X_i$ and $T = S_N$.

   (a) If, for example, a prior distribution G on Θ is constructed by taking M uniform on $\{1, 2, ..., N\}$ independent of μ_1 exponential with mean 1, independent of μ_2 exponential with mean 1, it is possible to explicitly find the (marginal) posterior of M given that $X_1 = x_1, ..., X_N = x_N$. Don't actually bother to finish the somewhat messy calculations needed to do this, but show that this is possible (indicate clearly why appropriate integrals can be evaluated explicitly).

   (b) Suppose now that G is constructed by taking M uniform on $\{1, 2, ..., N\}$ independent of (μ_1,μ_2) with joint density $g(\cdot, \cdot)$ on $(0, \infty) \times (0, \infty)$. Describe in as much detail as possible a "*Gibbs Sampling*" method for approximating the posterior of M, given $X_1 = x_1, ..., X_N = x_N$. (Give the necessary conditionals up to multiplicative constants, say how you're going to use them and what you'll do with any vectors you produce by simulation.)

Solution:

(a)

$$
\begin{aligned}
g(m, \lambda_1, \lambda_2 | x) &= \frac{\left(\prod_{i=1}^m e^{-\lambda_1}\frac{\lambda_1^{x_i}}{x_i!}\right)\left(\prod_{i=m+1}^N e^{-\lambda 2}\frac{\lambda_2^{x_i}}{x_i!}\right)\cdot \frac{1}{N} e^{-\lambda_1} e^{-\lambda 2}}{\sum_{m=1}^N \int_0^\infty \int_0^\infty \left(\prod_{j=1}^m e^{-\lambda_1}\frac{\lambda_1^{x_i}}{x_i!}\right)\left(\prod_{i=m+1}^N e^{-\lambda 2}\frac{\lambda_2^{x_i}}{x_i!}\right)\cdot \frac{1}{N} e^{-\lambda_1} e^{-\lambda 2} d\lambda_1 d\lambda_2} \\
&= \frac{e^{-(m+1)\lambda_1}\lambda_1^{S_m} e^{-(N-m+1)\lambda_2}\lambda_2^{S_N - S_m}}{\sum_{m=1}^N \int_0^\infty \int_0^\infty e^{-(m+1)\lambda_1}\lambda_1^{S_m} e^{-(N-m+1)\lambda_2}\lambda_2^{S_N - S_m} d\lambda_1 d\lambda_2}
\end{aligned}
$$

We can see that there are two *Gamma* integrals in the denominator, which can be calculated explicitly and so is the summation.
Then $\int_0^\infty \int_0^\infty g(m, \lambda_1, \lambda_2|x)\, d\lambda_1 d\lambda_2$ is actually the product of two *Gamma* integral multiplies some scalar, which surely can be evaluated explicitly.

(b) Let $\Theta = (m, \lambda_1, \lambda_2), \forall \theta \in \Theta$

$$g(\theta|x) = \frac{\lambda_1^{S_m} \cdot e^{-m\lambda_1} \cdot \lambda_2^{S_N - S_m} e^{-(N-m)\lambda_2} \cdot g(\lambda_1, \lambda_2)}{\sum_{m=1}^{M} \int\int \lambda_1^{S_m} \cdot e^{-m\lambda_1} \cdot \lambda_2^{S_N - S_m} e^{-(N-m)\lambda_2} \cdot g(\lambda_1, \lambda_2)\, d\lambda_1 d\lambda_2}$$

Gibbs Sampling algorithm:
(1) Generate $m_0 \sim U(1, ..., M)$, $(\lambda_{01}, \lambda_{02}) \sim g(\lambda_1, \lambda_2)$, and let $\theta_0 = (m_0, \lambda_{01}, \lambda_{02})$.
(2) With $\theta_i^* = \left(m_i^*, \lambda_{i,1}^*, \lambda_{i,2}^*\right)$ we generate $\theta_{i+1}^* = \left(m_i^*, \lambda_{i+1,1}^*, \lambda_{i+1,2}^*\right)$ in the following way:

- Generate m_{i+1}^* from $g\left(m|\lambda_{i,1}^*, \lambda_{i,2}^*, x\right)$. Since $g\left(m|\lambda_{i,1}^*, \lambda_{i,2}^*, x\right) \propto \left(\frac{\lambda_{i,1}^*}{\lambda_{i,2}^*}\right)^{S_m} \cdot e^{-m\left(\lambda_{i,1}^* - \lambda_{i,2}^*\right)}$.
  Use rejection algorithm: $h(M) = \frac{1}{M}$, $M_{m,i} = M \cdot \max_m \left(\left(\frac{\lambda_{i,1}^*}{\lambda_{i,2}^*}\right)^{S_m} \cdot e^{-m\left(\lambda_{i,1}^* - \lambda_{i,2}^*\right)}\right)$
  Hence $M_{m,i} h(M) > \left(\frac{\lambda_{i,1}^*}{\lambda_{i,2}^*}\right)^{S_m} \cdot e^{-m\left(\lambda_{i,1}^* - \lambda_{i,2}^*\right)}$. So we can use rejection algorithm to generate M_{i+1}^* from $g\left(m|\lambda_{i,1}^*, \lambda_{i,2}^*, x\right)$.
- Generate $\lambda_{i+1,1}^*$ from $g\left(\lambda_1|m_{i+1}^*, \lambda_{i,2}^*, x\right)$.
  Since $g\left(\lambda_1|m_{i+1}^*, \lambda_{i,2}^*, x\right) \propto \lambda_1^{S_{m_{i+1}^*}} \cdot e^{-m_{i+1}^* \lambda_1} g\left(\lambda_1|\lambda_{i,2}^*\right)$.
  $h(\lambda_1) = g\left(\lambda_1|\lambda_{i,2}^*\right)$.$M_{\lambda_1,i} = \max_{\lambda_1} \left(\lambda_1^{S_{m_{i+1}^*}} \cdot e^{-m_{i+1}^* \lambda_1}\right)$.
  So $M_{\lambda_1 \cdot i} \cdot h(\lambda_1) > \lambda_1^{S_{m_{i+1}^*}} \cdot e^{-m_{i+1}^* \lambda_1} g\left(\lambda_1|\lambda_{i,2}^*\right)$
  So we can generate $\lambda_{i+1,1}^*$ from $g\left(\lambda_1|m_{i+1}^*, \lambda_{i,2}^*, x\right)$.
- Generate $\lambda_{i+1,2}^*$ from $g\left(\lambda_2|m_{i+1}^*, \lambda_{i+1,1}^*, x\right)$.
  Since $g\left(\lambda_2|m_{i+1}^*, \lambda_{i+1,1}^*, x\right) \propto \lambda_2^{S_N - S_{m_{i+1}^*}} \cdot e^{-\left(N - m_{i+1}^*\right)\lambda_2} g\left(\lambda_2|\lambda_{i+1,1}^*\right)$
  $h(\lambda_2) = g\left(\lambda_2|\lambda_{i+1}^*, 1\right)$. $M_{\lambda_2,i} = \max_{\lambda_2} \left(\lambda_2^{S_N - S_{m_{i+1}^*}} \cdot e^{-\left(N - m_{i+1}^*\right)\lambda_2}\right)$.
  So $M_{\lambda_2,i} \cdot h(\lambda_2) > \lambda_2^{S_N - S_{m_{i+1}^*}} \cdot e^{-\left(N - m_{i+1}^*\right)\lambda_2} \cdot g\left(\lambda_2|\lambda_{i+1,1}^*\right)$.
  So we can get $\lambda_{i+1,2}^*$ from $g\left(\lambda_2|m_{i+1}^*, \lambda_{i+1,1}^*, x\right)$.
  Sequentially, we can get $\left(m_1^*, \lambda_{1,1}^*, \lambda_{1,2}^*\right), \cdots, \left(m_n^*, \lambda_{n,1}^*, \lambda_{n,2}^*\right), ...$
  We will use $m_1^*, m_2^*, \ldots, m_n^*, \ldots$ and the Ecdf is converge to $g(m|x)$.

4. Consider the simple two-dimensional discrete (posterior) distribution for $\theta = (\theta_1, \theta_2)$ given in the table below.

		θ_2			
		1	2	3	4
	4	0	0	.2	.1
	3	0	0	.05	.05
θ_1	2	.2	.1	0	0
	1	.2	.1	0	0

(a) This is obviously a very simple set-up where anything one might wish to know or say about the distribution is easily derivable from the table. However, for sake of example, suppose that one wished to use SSS to approximate $Q = P[\theta 1 \leq 2] = .6$. Argue carefully that *Gibbs Sampling* will fail here to produce a correct estimate of Q. In qualitative terms, what is it about this joint distribution that causes this failure?

(b) Consider the very simple version of the Metropolis-Hastings algorithm where for each i, $J_i(\theta_0|\theta)$ is uniform on the 8 vectors θ_0 where the posterior probability in the table above is positive. Make out an 8? table giving the conditional probabilities that (under the *Metropolis-Hastings* algorithm) $\theta_i^* = \theta'$ given $\theta_{i-1}^* = \theta$. If infact θ_{i-1}^* has the distribution in the table, show that θ_i^* also has this distribution. (This is the important fact that the algorithm has the target distribution as its "*stationary distribution.*")

Solution:

(a) We may compute $P(\theta_1|\theta_2)$ and $P(\theta_2|\theta_1)$.

		θ_1			
		1	2	3	4
	4	0	0	2/3	2/3
	3	0	0	.5	.5
θ_2	2	2/3	1/3	0	0
	1	2/3	1/3	0	0

$P(\theta_1|\theta_2)$

		θ_1			
		1	2	3	4
	4	0	0	.8	2/3
	3	0	0	.2	1/3
θ_2	2	.5	.5	0	0
	1	.5	.5	0	0

$P(\theta_2|\theta_1)$

The state space is transient, that is the state 1 is communicate only to state 2 and state 3 is communicate only to state 4. So if we use *Gibbs* sampling to produce the simulated values, WLOG, assume we generate θ_1^* first and then generate θ_2^* based on $P(\theta_2|\theta_1^*)$ value, then if θ_1 was in state 1 or 2 then θ_2 can never get to state 3 or 4, vice versa. Then $p(\theta_1 \leq 2)$ is either 1 or 0. Hence, the *Gibbs* Sampling does not work in this case.

(b) Look at the following table·

<table>
<tr><th>θ</th><th>θ'</th><th>(1,1)</th><th>(1,2)</th><th>(2,1)</th><th>(2,2)</th><th>(3,3)</th><th>(3,4)</th><th>(4,3)</th><th>(4,4)</th></tr>
<tr><td>.2</td><td>θ_1 (1,1)</td><td colspan="8" rowspan="8">1/8</td></tr>
<tr><td>-1</td><td>θ_2 (1,2)</td></tr>
<tr><td>.2</td><td>θ_3 (2,1)</td></tr>
<tr><td>-1</td><td>θ_4 (2,2)</td></tr>
<tr><td>.05</td><td>θ_5 (3,3)</td></tr>
<tr><td>.05</td><td>θ_6 (3,4)</td></tr>
<tr><td>.2</td><td>θ_7 (4,3)</td></tr>
<tr><td>.1</td><td>θ_8 (4,4)</td></tr>
</table>

I will specify $J\left(\theta'|\theta\right) = \frac{1}{8}11'$. i.e. all jump probability equals to 1/8.
Under these $J\left(\theta'|\theta\right)$

$$\left(P\left(\theta_i^* = \theta_j|\theta_{i-1}^* = \theta_i\right)\right)_{i,j} = \frac{1}{8}\begin{pmatrix} \theta_1 & \theta_2 & \theta_3 & \theta_4 & \theta_5 & \theta_6 & \theta_7 & \theta_8 \\ 4 & .5 & 1 & .5 & .25 & .25 & 1 & .5 \\ 1 & 2 & 1 & 1 & .5 & .5 & 1 & 1 \\ 1 & .5 & 4 & .5 & .25 & .25 & 1 & .5 \\ 1 & 1 & 1 & 2 & .5 & .5 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & .5 & 1 & .5 & .25 & .25 & 4 & .5 \\ 1 & 1 & 1 & 1 & .5 & .25 & 1 & 2 \end{pmatrix}\begin{pmatrix} \\ \theta_1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \\ \theta_5 \\ \theta_6 \\ \theta_7 \\ \theta_8 \end{pmatrix}$$

denote the matrix $\frac{1}{8}T$.
If $\theta_{i-1}^* \sim P$ then

$$\begin{aligned} P\left(\theta_i^* = \theta\right) &= P\left(\theta_{i-1}^* = \theta_i\right) \cdot P\left(\theta_i^* = \theta_j|\theta_{i-1}^* = \theta_i\right) \\ &= (.2, .1, .2, .1, .05, .05, .2, .1) \cdot \frac{1}{8}T \\ &= (.2, .1, .2, .1, .05, .05, .2, .1) \end{aligned}$$

So, $\theta_i^* \sim P$.

5. Suppose that$X_1, X_2, ..., X_n$ are iid $N(\theta, 1)$ and that $\gamma(\theta) = E_\theta X_1^2 = \theta^2 + 1$ is of interest. In class *Vardeman Rao-Blackwellized* a method of moments estimator of $\gamma(\theta)$ and obtained the estimator $\delta^*(X) = 1 + \overline{X}^2 - \frac{1}{n}$. Argue that this (silly) estimator of $\gamma(\theta)$ is the UMVUE. Argue carefully that the estimator

$$\delta^{**}(X) = \begin{cases} \delta^*(X) & \text{if } \overline{X}^2 > \frac{1}{n} \\ 1 & \text{if } \overline{X}^2 \leq \frac{1}{n} \end{cases}$$

is strictly better than $\delta^*(X)$ (that is, $\text{MSE}_\theta(\sigma^{**}(X)) < \text{MSE}_\theta(\sigma^{**}(X)\ \forall\theta.)$ (Hint: What can you say about the random variable $(\delta^*(X) - \gamma(\theta))^2 - (\delta^{**}(X) - \gamma(\theta))^2$?) Is $\delta^{**}(X)$

unbiased?

Solution:

As $E_\theta X_1^2 = \theta^2 + 1$, $E_\theta \bar{X}^2 = \text{var}\left(\bar{X}\right) + \left(E\bar{X}\right)^2 = \frac{1}{n} + \theta^2$, then
$E\left(\delta^*(x)\right) = E\left(1 + \bar{X}^2 - \frac{1}{n}\right) = 1 + \theta^2 \Longrightarrow \delta^*(x)$ is unbiased est. of $\gamma(\theta)$.
As x is exponential family and the natural parameter space contains interior.
Thus $\bar{X}$ is a sufficient & complete statistics.
By *Lehman-Scheffė* theorem, $\delta^*(x)$ is UMVUE.

$$\begin{aligned} & E\left(\delta^*(x) - \gamma(\theta)\right)^2 - \left(\delta^{**}(x) - \gamma(\theta)\right)^2 \\ = & E\left(E\left\{\left(\delta^*(x) - \gamma(\theta)\right)^2 - \left(\delta^{**}(x) - \gamma(\theta)\right)^2 |\bar{X}\right\}\right) \\ = & E\left\{0 \cdot I\left(\bar{X}^2 > \frac{1}{n}\right) + \left[\left(\delta^*(x) - \gamma(\theta)\right)^2 - \left(1 - \gamma(\theta)\right)^2\right] |I\left(\bar{X}^2 \leq \frac{1}{n}\right)\right\} < 0 \\ \text{as } \delta^*(x) < & 1 \text{ when } \bar{x}^2 < \frac{1}{n} \text{ and } P\left(\bar{x}^2 < \frac{1}{n}\right) > 0 \end{aligned}$$

$.\Longrightarrow E\left(\delta^*(x) - \gamma(\theta)\right)^2 < E\left(\delta^{**}(x) - \gamma(\theta)\right)^2 \Longrightarrow MSE_\theta\left(\delta^*(\theta)\right) < MSE_\theta\left(\delta^{**}(\theta)\right)$
Then $\delta^{**}(\theta)$ is strictly better than $\delta^*(\theta)$.

6. (Gartwaithe, Jolliffe and Jones) Suppose that $X_1, X_2, ..., X_n$ are iid with marginal *pdf*

$$f(x|\theta) = I\left[1 \leq x \leq \theta\right]\theta x^{\theta-1}$$

for $\theta > 0$. $-\ln(X_1)$ is an unbiased estimator of θ. Find a better one (one with smaller variance). Is your estimator UMVU? Explain.

Solution:

$f(x) = \theta \cdot x^{-1+\theta}$ Let $y = -\ln x$, $\Rightarrow f(y) = \theta \cdot \exp(-y\theta) \sim \exp(1/\theta)$
$E(-\ln x_1) = \frac{1}{\theta} \Longrightarrow -\ln x_1$ is unbiased.
As $f(\underline{x}) = \exp\left\{-(1+\theta)\sum \ln x_i + n\ln\theta\right\}$ then x belongs exponential family and natural parameter space contains interior, we have $-\sum_{i=1}^n \ln x_i$ is sufficient and complete.
And also $E\left(-\frac{\sum \ln x_i}{n}\right) = \theta^{-1}$ By *Lehmam-Scheffė* theorem, $-\frac{\sum \ln x_i}{n}$ is UNVUE.
This estimator is better than $-\ln x_1$ as $P\left(-\frac{\sum \ln x_i}{n} \neq -\ln x_1\right) > 0$.

7. As an example of an ad hoc (non-exponential-family) application of the *Lehmann-Scheffė* argument, consider the following. Suppose that X_1, X_2 are iid with marginal pmf on the positive integers

$$f(x|\theta) = \frac{1}{\theta}I\left[1 \leq x \leq \theta\right]$$

for θ a positive integer.

(a) Show that the statistic $Y = \max(X_1, X_2)$ is both sufficient and complete. (Look on the handouts for Bahadur's Theorem and the *Lehmann-Scheffė* Theoremfor a definition of completeness.) (Think about this problem in geometric terms (on the grid of integer pairs in the (x_1, x_2)-plane) in order to work out the distribution of Y .)

(b) Then argue that

$$\delta(X) = \frac{Y^3 - (Y-1)^3}{Y^2 - (Y-1)^2}$$

is a UMVUE of θ.

(c) Note that $I[X_1 = 1]$ is an unbiased estimator of θ^{-1}. *Rao-Blackwellize* this using Y . Is your resulting function of Y a UMVUE of θ^{-1}?

Solution:

(a) $f(x_1, x_2|\theta) = \frac{1}{\theta^2} I\{1 \leq \min(x_1, x_2)\} I\{\max(x_1, x_2) \leq \theta\}$. By factorization theorem, $Y = \max(X_1, X_2)$ is sufficient for θ.
Suppose $E_\theta h(\max(x_1, x_2)) = 0, \forall\theta. \Rightarrow h(x_1) P(x_1 \geq x_2) + h(x_2) P(x_1 < x_2) = 0$.
for $\forall\theta > 1$, $P(x_1 \geq x_2) = \frac{\theta+1}{2\theta} > 0$.
$P(x_1 < x_2) = \frac{\theta-1}{2\theta} > 0$.
$\Rightarrow h(x_1) = 0 \;\&\; h(x_2) = 0 \;\forall\theta > 1$
for $\theta = 1$, $h(x) P_\theta(x_1 = x_2 = 1) = 0$
$\Rightarrow h(x_1) = h(x_2) = 0$.
Then $h(\max(x_1, x_2)) = 0.$ w.p.1 $\forall\theta$
i.e. $P_\theta(h(\max(x_1, x_2)) = 0) = 1 \quad \Rightarrow Y = \max(X_1, X_2)$. is complete.

(b) $P(\max(X_1, X_2) = t) = P(X_1 \leq t, X_2 \leq t) - P(X_1 \leq t-1, X_2 \leq t-1)$
$= \frac{t^2}{\theta^2} - \frac{(t-1)^2}{\theta^2} = \frac{2t-1}{\theta^2}$.
Consider $\hat{\theta} = 2X_1 - 1$, which is an unbiased estimator.
$E(2X_1 - 1|Y = \max(X_1, X_2) = t) = 2E(X_1|Y = t) - 1$
$= 2\left\{t \cdot P(X_1 = t|Y = t) + \sum_{k=1}^{t-1} kP(X_1 = k|Y = t)\right\} - 1$
$= 2\left\{t \cdot \frac{P(X_1=t, X_2 \leq t)}{P(Y=t)} + \sum_{k=1}^{t-1} \frac{kP(X_1=k, X_2=t)}{P(Y=t)}\right\} - 1$
$= 2\left\{t \cdot \frac{t/\theta^2}{(2t-1)/\theta^2} + \frac{\frac{t(t-1)}{2} \cdot \frac{1}{\theta^2}}{\frac{(2t-1)}{\theta^2}}\right\} - 1 = 2\left\{\frac{t^2}{2t-1} + \frac{t(t-1)}{2(2t-1)}\right\} - 1 = \frac{3t^2-3t+1}{2t-1}$.
$\Rightarrow E(2X_1 - 1|Y = \max(X_1, X_2)) = \frac{3Y^2-3Y+1}{2Y-1} = \frac{Y^3-(Y-1)^3}{Y^2-(Y-1)^2} = g(Y)$,By *Rao-Blackwell* theorem, $g(Y)$ is UMVUE.

(c) $E(I(X_1 = 1)|Y = \max(X_1, X)) = \frac{1}{2Y-1}$. As I$(X_1 = 1)$ is an unbiased estimator of θ^{-1}, by *Rao-Blackwell* theorem, $\frac{1}{2Y-1}$ is a UMVUE of θ^{-1}.

8. (Knight) **Optional** (not required but recommended) Suppose that $X_1, X_2, ..., X_n$ are iid exponential with mean β, i.e. with marginal density

$$f(x|\beta) = I[0 < x]\frac{1}{\beta}\exp\left(-\frac{x}{\beta}\right)$$

Let

$$\gamma(\beta) = \exp\left(-\frac{t}{\beta}\right)$$

for some fixed number, t.($\gamma(\beta)$ is the probability that $X_1 > t$.)

(a) Show that for every $\beta, \sum_{i=1}^n X_i$ and $X_1/\sum_{i=1}^n X_i$ are independent random variables.

(b) *Rao-Blackwellize* $I[X1 > t]$ using the natural sufficient statistic here. Is the result a UMVUE of $\gamma(\beta)$? Explain.

Solution:

(a) $X \sim \exp(\beta)$. $\Rightarrow \sum_1^n X_i \sim \text{Gamma}(n, \beta) \Rightarrow \sum_2^n X_i \sim \text{Gamma}(n-1, \beta)$

Also $X_1/\sum_1^n X_i \sim \text{Beta}(1, n-1)$, which is an ancillary statistic. And $\sum_1^n X_i$ is a sufficient and complete statistic.

By *Basu's Theorem* $\sum X_i$ and $X_1/\sum_1^n X_i$ are independent.

Details for showing $X_1/\sum_1^n X_i \sim \text{Beta}(1, n-1)$:

Let $Y = X_1/\sum_1^n X_i, X = X_1$

$$\begin{cases} X_1 = X \\ \sum_2^n X_i = \frac{X}{Y} - X \end{cases}, \quad \frac{\partial\left(X_1, \sum_2^n X_i\right)}{\partial(X,Y)} = -\frac{X}{Y^2}$$

Then

$$f_{X,Y}(x,y) = f_{X_1,\sum_2^n X_i}\left(x, \frac{x}{y} - x\right)\cdot\frac{x}{y^2}$$

$$= \frac{1}{\beta}e^{-x/\beta}\frac{(x/y-x)^{n-2}e^{-\frac{x(1-y)}{y^2\beta}}}{\Gamma(n-1)\beta^{n-1}}\cdot\frac{x}{y^2} = \frac{1}{\Gamma(n-1)\beta^n}\frac{x^{n-1}(1-y)^{n-2}}{y^n}e^{-\frac{x}{y\beta}}$$

$$\Rightarrow f(y) = \int_0^\infty \frac{1}{\Gamma(n-1)\beta^n}\frac{x^{n-1}}{y^n}(1-y)^{n-2}e^{-\frac{x}{y\beta}}dx$$

$$= \frac{1}{\Gamma(n-1)\beta^n}\int_0^\infty t^{n-1}(1-y)^{n-2}e^{-t/\beta}dt = (n-1)(1-y)^{n-2}$$

$$\Rightarrow Y \sim \text{Beta}(1, n-1)$$

(b) By exponential family, we have $\sum X_i$ is complete and sufficient for θ .

$$\delta(T) = E\left\{I(x_1 > t)\,|\sum x_i\right\} = P(x_1 > t|\sum x_i) = P\left(\frac{x_1}{\sum x_i} > \frac{t}{\sum x_i}|\sum x_i = T\right)$$

As $\frac{x_1}{\sum x_i}$ follows $\text{Beta}(1, n-1)$ dsn

$$= \int_{\frac{t}{T}}^1 \frac{1}{n-1}\cdot(1-x)^{n-2}dx = \left(1 - \frac{t}{T}\right)^{n-1}.$$

$$\Rightarrow \delta(T) = \left(1 - \frac{t}{T}\right)^{n-1} \text{ is UMVUE for } \gamma(\beta) = \exp\left(-\frac{t}{\beta}\right).$$

As $I(x_1 > t)$ is an unbiased estimator of $\gamma(\beta)$, and T is complete & sufficient statistics, then by *Lehmann-Scheffé* theorem $\delta(T) = \left(1 - \frac{t}{T}\right)^{n-1}$ is UMVUE.