CS 661, Spring 2014
Midterm Solutions

1. (a) We first need to compute the density f of F by differentiating F to get

$$f(x) = \begin{cases} 2\lambda x \exp(-\lambda x^2) & \text{if } x \geq 0, \\ 0 & \text{if } x < 0, \end{cases}$$

Then the likelihood function $L_n(\alpha)$ of an IID set of samples $X_1, \ldots, X_n$ is the product of the density function evaluated at the n observations, so

$$L_n(\lambda) = \prod_{i=1}^{n} f(X_i) = (2\lambda)^n \left(\prod_{i=1}^{n} X_i \right) \exp\left\{ \sum_{j=1}^{n} -\lambda X_j^2 \right\}.$$

Taking the natural log gives the log-likelihood function

$$\mathcal{L}_n(\lambda) = \ln(L_n(\lambda)) = n \ln 2 + n \ln \lambda + \sum_{i=1}^{n} \ln X_i - \lambda \sum_{j=1}^{n} X_j^2.$$

We then compute the derivative of $\mathcal{L}_n(\lambda)$ with respect to λ to get

$$\frac{d}{d\lambda} \mathcal{L}_n(\lambda) = \frac{n}{\lambda} - \sum_{j=1}^{n} X_j^2.$$

Then equating this to 0 and solving for α yields

$$\hat{\lambda} = \frac{n}{\sum_{j=1}^{n} X_j^2}.$$

The second derivative of $\mathcal{L}_n(\lambda)$ is

$$\frac{d^2}{d\lambda^2} \mathcal{L}_n(\lambda) = -\frac{n}{\lambda^2} < 0,$$

so $\hat{\lambda}$ is a local maximizer of $\mathcal{L}_n(\lambda)$ and $L_n(\lambda)$. Evaluating $\mathcal{L}_n(\lambda)$ at the boundary $\lambda = 0$ is $-\infty$, so $\hat{\alpha}$ is the global maximizer of $\mathcal{L}_n(\alpha)$ and $L_n(\alpha)$. Thus, $\hat{\alpha}$ is the MLE of α.

(b) For a gamma(α, β) distribution, the population mean is $\mu_1 = \alpha\beta$ and the population variance is $\alpha\beta^2$, so the population second moment is $\mu_2 = \alpha\beta^2 + \alpha^2\beta^2$. Let $M_k = \frac{1}{n} \sum_{i=1}^{n} X_i^k$ be the kth sample moment, $k = 1, 2$. Then set $\mu_1 = M_1$ and $\mu_2 = M_2$, giving us the equations $\alpha\beta = M_1$ and $\alpha\beta^2 + \alpha^2\beta^2 = M_2$, and we want to solve for α and β in these two equations. The first equation yields $\alpha = M_1/\beta$, and substituting this into the second equation gives $M_1\beta + M_1^2 = M_2$, so $\beta = (M_2 - M_1^2)/M_1$. Then substituting this into the first equation gives $\alpha = M_1^2/(M_2 - M_1^2)$. Thus, the method-of-moments estimators of α and β are $\hat{\alpha} = M_1^2/(M_2 - M_1^2)$ and $\hat{\beta} = (M_2 - M_1^2)/M_1$.

2. (a) For each $i = 1, 2, \ldots, n$, let $W_i = I(2 \le 5X_i + \ln Y_i \le 9)$, where I denotes the indicator function, which equals 1 (resp., 0) when its argument is true (resp., false). We define $\hat{\theta}_n$ as

$$\hat{\theta}_n = \frac{1}{n} \sum_{i=1}^{n} W_i$$

(b) Note that

$$\mathcal{E}[\hat{\theta}_n] = \frac{1}{n} \sum_{i=1}^{n} \mathcal{E}[I(2 \le 5X_i + \ln Y_i \le 9)] = \frac{1}{n} \sum_{i=1}^{n} \mathcal{P}(2 \le 5X_i + \ln Y_i \le 9) = \theta$$

since the (X_i, Y_i) are IID pairs. Hence, $\hat{\theta}_n$ is an unbiased estimator of θ.

(c) Note that W_i, $i = 1, 2, \ldots, n$, are IID because (X_i, Y_i), $i = 1, 2, \ldots, n$, are IID pairs. Also, $\hat{\theta}_n$ is the sample average of the W_i, $i = 1, 2, \ldots, n$, each of which has mean $\mathcal{E}[W_i] = \theta$, which we assumed satisfies $0 < \theta < 1$, so θ is finite. Hence, $\hat{\theta}_n \approx \theta$ for large n by the ordinary law of large numbers.

(d) Since W_i is binary, $W_i^2 = W_i$, so we have $\beta^2 = \mathcal{V}[W_i] = \mathcal{E}[W_i^2] - (\mathcal{E}[W_i])^2 = \mathcal{E}[W_i] - (\mathcal{E}[W_i])^2 = \theta(1-\theta)$, which is positive and finite since we assumed $0 < \theta < 1$. Also, since $\hat{\theta}_n$ is the sample average of the W_i, which are IID with mean θ, we have the following CLT:

$$\frac{\sqrt{n}}{\sqrt{\theta(1-\theta)}} (\hat{\theta}_n - \theta) \stackrel{\mathcal{D}}{\approx} \mathcal{N}(0, 1)$$

for large n. Thus, $\beta = \sqrt{\theta(1-\theta)}$ and $\alpha = \theta$.

(e) An estimator for β is $\hat{\beta}_n = \sqrt{\hat{\theta}_n(1-\hat{\theta}_n)}$. Another possible estimator of β is

$$\hat{\beta}_n = \left[\frac{1}{n-1} \sum_{i=1}^{n} (W_i - \hat{\theta}_n)^2 \right]^{1/2}.$$

Since $\hat{\theta}_n \approx \theta$ for large n, as we argued in part (c), the first estimator satisfies $\hat{\beta}_n = \sqrt{\hat{\theta}_n(1-\hat{\theta}_n)} \approx \sqrt{\theta(1-\theta)} = \beta$ for large n, so it satisfies a law of large numbers. The second estimator also does since it is the sample variance of the $I(X_i \in A)$, $i = 1, 2, \ldots, n$, and we know that the sample variance of IID samples is consistent, as can be seen as follows. Note that the second estimator satisfies

$$\hat{\beta}_n = \left[\frac{n}{n-1} \frac{1}{n} \sum_{i=1}^{n} W_i - \frac{n}{n-1} \hat{\theta}_n^2 \right]^{1/2}.$$

Now $n/(n-1) \approx 1$ for large n, and $\frac{1}{n} \sum_{i=1}^{n} W_i \approx \mathcal{E}[W_i] = \theta$ by the law of large numbers. Thus, $\hat{\beta}_n \approx \sqrt{\theta - \theta^2} = \beta$.

(f) The CLT implies that for large n,

$$\begin{aligned}
0.95 &= \mathcal{P}\left\{\mathcal{N}(0,1) < 1.65\right\} \\
&\approx \mathcal{P}\left\{\frac{\sqrt{n}}{\beta}(\hat{\theta}_n - \theta) < 1.65\right\} \\
&\approx \mathcal{P}\left\{\frac{\sqrt{n}}{\hat{\beta}_n}(\hat{\theta}_n - \theta) < 1.65\right\} \\
&= \mathcal{P}\left\{\hat{\theta}_n - \theta < 1.65\frac{\hat{\beta}_n}{\sqrt{n}}\right\} \\
&= \mathcal{P}\left\{\hat{\theta}_n - 1.65\frac{\hat{\beta}_n}{\sqrt{n}} < \theta\right\},
\end{aligned}$$

where the third step holds since $\hat{\beta}_n \approx \beta$ for large n for $\hat{\beta}_n$ defined as either of the estimators in the previous part. Thus, an approximate 95% lower-one-sided confidence interval for θ is

$$\left(\hat{\theta}_n - 1.65\frac{\hat{\beta}_n}{\sqrt{n}},\, +\infty\right)$$

Since we θ is a probability, we know that $\theta \leq 1$, so we can replace the upper confidence limit by 1 to get

$$\left(\hat{\theta}_n - 1.65\frac{\hat{\beta}_n}{\sqrt{n}},\, 1\right]$$

as our approximate 95% lower-one-sided confidence interval for θ.

3. (a) Since f_y is the density of a lognormal(μ, y^2) random variable, we can generate a random variate with this density by first generating a $\mathcal{N}(\mu, y^2)$ random variate and then exponentiating it. Thus, first generate $U \sim \text{unif}[0,1]$, and the compute $V = \Phi^{-1}(U)$, where Φ^{-1} is the inverse of the standard normal CDF, which we compute using the black box. Then let $W = \mu + yU$, which has $\mathcal{N}(\mu, y^2)$ distribution. Finally, let $X = e^W$. So the complete algorithm is

> 1. Generate $U \sim \text{unif}[0,1]$.
> 2. Return $X = \exp\{\mu + y\Phi^{-1}(U)\}$.

(b) We first compute the CDF G corresponding to density g. For $3 \leq x < 4$, we have

$$G(x) = \int_3^x \frac{t-3}{2}dt = \left[\frac{t^2}{4} - \frac{3t}{2}\right]_3^x = \frac{x^2}{4} - \frac{3x}{2} + \frac{9}{4},$$

so $G(4) = 1/4$. For $4 < x \leq 7$, we have

$$G(x) = \int_3^4 \frac{t-3}{2} dt + \int_4^x \frac{7-t}{6} dt = \frac{1}{4} + \left[\frac{7t}{6} - \frac{t^2}{12}\right]_4^x = \frac{7x}{6} - \frac{x^2}{12} - \frac{37}{12}.$$

Putting this together we have that

$$G(x) = \begin{cases} 0 & \text{if } x < 3, \\ (x^2 - 6x + 9)/4 & \text{if } 3 \leq x \leq 4, \\ (-x^2 + 14x - 37)/12 & \text{if } 4 < x \leq 7, \\ 1 & \text{if } x > 7. \end{cases}$$

Thus, to generate a random variate $Y \sim G$, we first generate $U \sim \text{unif}[0, 1]$, and if $0 \leq U \leq 1/4$, we invert the second line of G; otherwise, we invert the third line of G.

In the first case (when $0 \leq U \leq 1/4$), set $(Y^2 - 6Y + 9)/4 = U$ and solve for Y in terms of U. Rearranging terms leads to $Y^2 - 6Y + 9 - 4U = 0$, and applying the quadratic formula gives $Y = (6 \pm \sqrt{36 - 36 + 4U})/2 = 3 \pm 2\sqrt{U}$. Since we should have $3 \leq Y \leq 4$ in this case, we take the positive root, so we set $Y = 3 + 2\sqrt{U}$.

For the second case (when $4 < x \leq 7$), set $(-Y^2 + 14Y - 37)/12 = U$, and solve for Y in terms of U. Rearranging gives $-Y^2 + 14Y - (37 + 12U) = 0$, so the quadratic equation yields $Y = (-14 \pm \sqrt{196 - 148 - 48U})/(-2) = 7 \pm 2\sqrt{3(1-U)}$. Since we should have $4 < Y \leq 7$, we take the negative root, so we set $Y = 7 - 2\sqrt{3(1-U)}$.

Putting this all together gives the following algorithm for generating Y with density g:

1. Generate $U \sim \text{unif}[0, 1]$.
2. If $0 \leq U \leq 1/4$, then return $Y = 3 + 2\sqrt{U}$ and stop.
3. If $1/4 < U \leq 1$, then return $Y = 7 - 2\sqrt{3(1-U)}$ and stop.

(c) Note that h is a continuous mixture of f_y and g, so we can generate a random variate Z by first generating $Y \sim g$, and then using Y^2 as the variance of normal with mean μ, which we exponentiate to get Z.

1. Generate $U_1 \sim \text{unif}[0, 1]$.
2. If $0 \leq U_1 \leq 1/4$, then set $Y = 3 + 2\sqrt{U_1}$.
3. If $1/4 < U_1 \leq 1$, then set $Y = 7 - 2\sqrt{3(1-U_1)}$.
4. Generate $U_2 \sim \text{unif}[0, 1]$, independent of U_1.
5. Return $Z = \exp\{\mu + Y\Phi^{-1}(U_2)\}$.

4. Note that $\ln Z = \sum_{i=1}^{n} \ln X_i$. When n is large, $\sum_{i=1}^{n} \ln X_i$ is the sum of a large number of i.i.d. random variables, which is approximately normally distributed by the central limit theorem. Also, $Z = \exp\{\ln Z\} = \exp\{\sum_{i=1}^{n} \ln X_i\}$, and we know that if Y is normally distributed, then e^Y has a lognormal distribution. Thus, the distribution of Z should be approximately lognormal when n is large.