STA 623 — Fall 2009 — Dr. Charnigo

Written Assignment 4 Solutions

1a. The cumulative distribution function is

$$F(x) := \int_{-\infty}^{x} \frac{\exp[-t]}{(1+\exp[-t])^2} dt = \frac{1}{1+\exp[-x]}.$$

Setting $F(x)$ equal to 0.5 yields $\exp[-x] = 1$ or $x = 0$, so the median is 0. To find the mean, write

$$\int_{-\infty}^{\infty} xf(x)\ dx = \int_{-\infty}^{0} xf(x)\ dx + \int_{0}^{\infty} xf(x)\ dx.$$

For $\int_{-\infty}^{0} xf(x)\ dx$, we integrate by parts with $u := x$, $dv := f(x)\ dx$, $v := F(x)$, and $du := dx$ to obtain

$$\begin{aligned}
\int_{-\infty}^{0} xf(x)\ dx &= xF(x)|_{-\infty}^{0} - \int_{-\infty}^{0} F(x)\ dx \\
&= \frac{x}{1+\exp[-x]}|_{-\infty}^{0} - \int_{-\infty}^{0} \frac{1}{1+\exp[-x]}\ dx \\
&= 0 - \int_{-\infty}^{0} \left(1 + \frac{-\exp[-x]}{1+\exp[-x]}\right)\ dx \\
&= -\lim_{M\to\infty} \left(x + \log[1+\exp[-x]]\right)|_{-M}^{0} \\
&= -\lim_{M\to\infty} (M + \log 2 - \log[1+\exp[M]]) \\
&= -\lim_{M\to\infty} (M + \log 2 - \log[\exp[M](\exp[-M]+1)]) \\
&= -\lim_{M\to\infty} (M + \log 2 - M - \log[\exp[-M]+1]) = -\log 2.
\end{aligned}$$

For $\int_{0}^{\infty} xf(x)\ dx$, we integrate by parts with $u := x$, $dv := f(x)\ dx$, $v := F(x) - 1$, and $du := dx$ to obtain

$$\begin{aligned}
\int_{0}^{\infty} xf(x)\ dx &= x[F(x)-1]|_{0}^{\infty} - \int_{0}^{\infty} [F(x)-1]\ dx \\
&= \frac{-x\exp[-x]}{1+\exp[-x]}|_{0}^{\infty} - \int_{0}^{\infty} \frac{-\exp[-x]}{1+\exp[-x]}\ dx \\
&= 0 - \log[1+\exp[-x]]|_{0}^{\infty} \\
&= \log 2.
\end{aligned}$$

So

$$\int_{-\infty}^{\infty} xf(x)\ dx = -\log 2 + \log 2 = 0.$$

An alternative (and easier) computation of the mean entails writing $f(x) = 1/(\exp[-x/2] + \exp[x/2])^2$, which makes clear that $xf(x)$ is an odd function. Hence $\int_{-\infty}^{\infty} xf(x)\ dx$ must be 0, since the integral is absolutely convergent: note that $|xf(x)| \leq |x|\exp[-|x|]$.

1b. Put

$$f(x;\mu,\sigma) := \frac{1}{\sigma} f\left(\frac{x-\mu}{\sigma}\right) = \frac{\exp[-(x-\mu)/\sigma]}{\sigma(1+\exp[-(x-\mu)/\sigma])^2}$$

for $\mu \in (-\infty,\infty)$ and $\sigma \in (0,\infty)$. Then $\{f(x;\mu,\sigma) : \mu \in (-\infty,\infty), \sigma \in (0,\infty)\}$ is a location-scale family.

1c. If $X \sim f(x;\mu,\sigma)$, then $Z := (X-\mu)/\sigma \sim f(z)$. Solving

$$P(Z \leq 0) = 0.5 = P(X \leq x_{0.5}) = P(Z \leq (x_{0.5}-\mu)/\sigma)$$

for $x_{0.5}$ yields $x_{0.5} = \mu$. Also, solving

$$(E[X] - \mu)/\sigma = E[(X - \mu)/\sigma] = E[Z] = 0$$

for $E[X]$ yields $E[X] = \mu$.

2a. We have

$$E[(1-X)g'(X)] = \int_0^1 (1-x)g'(x)\frac{\Gamma[\alpha+\beta]}{\Gamma[\alpha]\Gamma[\beta]}x^{\alpha-1}(1-x)^{\beta-1}\,dx = \int_0^1 \frac{\Gamma[\alpha+\beta]}{\Gamma[\alpha]\Gamma[\beta]}x^{\alpha-1}(1-x)^{\beta}g'(x)\,dx.$$

Integrating by parts with $u := \frac{\Gamma[\alpha+\beta]}{\Gamma[\alpha]\Gamma[\beta]}x^{\alpha-1}(1-x)^{\beta}$, $dv := g'(x)\,dx$, $v := g(x)$, and $du := \frac{\Gamma[\alpha+\beta]}{\Gamma[\alpha]\Gamma[\beta]}x^{\alpha-2}(1-x)^{\beta-1}((\alpha-1)(1-x)-\beta x)\,dx$ yields

$$\begin{aligned}
E[(1-X)g'(X)] &= \left(\frac{\Gamma[\alpha+\beta]}{\Gamma[\alpha]\Gamma[\beta]}x^{\alpha-1}(1-x)^{\beta}g(x)\right)|_0^1 - \int_0^1 g(x)\frac{\Gamma[\alpha+\beta]}{\Gamma[\alpha]\Gamma[\beta]}x^{\alpha-2}(1-x)^{\beta-1}((\alpha-1)(1-x)-\beta x)\,dx \\
&= 0 - \int_0^1 g(x)((\alpha-1)(1-x)/x - \beta)\frac{\Gamma[\alpha+\beta]}{\Gamma[\alpha]\Gamma[\beta]}x^{\alpha-1}(1-x)^{\beta-1}\,dx \\
&= \int_0^1 g(x)(\beta - (\alpha-1)(1-x)/x)f(x;\alpha,\beta)\,dx \\
&= E[g(X)(\beta - (\alpha-1)(1-X)/X)].
\end{aligned}$$

2b. Exercise 3b of the take-home midterm examination showed that

$$E[X^{\gamma}(1-X)^{\delta}] = \frac{\Gamma[\alpha+\beta]\Gamma[\alpha+\gamma]\Gamma[\beta+\delta]}{\Gamma[\alpha]\Gamma[\beta]\Gamma[\alpha+\beta+\gamma+\delta]}$$

whenever $\gamma, \delta \geq 0$, a result that is obviously still valid for negative γ or δ provided that $\alpha+\gamma$ and $\beta+\delta$ are positive. Since $\lim_{x\to 0}\log x\ x^C = \lim_{x\to 1}\log x\ (1-x)^C = 0$ for any positive integer C, putting $C := \beta = (\alpha - 1)$ and $g(x) := \log x$ yields

$$\begin{aligned}
CE[\log X(2 - X^{-1})] &= E[\log X(C - C(1-X)/X)] \\
&= E[(1-X)/X] \\
&= E[X^{-1}(1-X)^1] \\
&= \frac{\Gamma[2C+1]\Gamma[C]\Gamma[C+1]}{\Gamma[C+1]\Gamma[C]\Gamma[2C+1]} \\
&= 1.
\end{aligned}$$

Therefore, $E[\log X(2 - X^{-1})] = 1/C$.

2c. We have

$$f(x;\alpha,\beta) = 1_{\{x\in(0,1)\}}\frac{\Gamma[\alpha+\beta]}{\Gamma[\alpha]\Gamma[\beta]}x^{\alpha-1}(1-x)^{\beta-1} = h(x)c(\theta)\exp[w_1(\theta)t_1(x) + w_2(\theta)t_2(x)]$$

with $h(x) := 1_{\{x\in(0,1)\}}$, $\theta := (\alpha,\beta)'$, $c(\theta) := \frac{\Gamma[\alpha+\beta]}{\Gamma[\alpha]\Gamma[\beta]}$, $w_1(\theta) := \alpha - 1$, $t_1(x) := \log x$, $w_2(\theta) := \beta - 1$, and $t_2(x) := \log(1-x)$. Thus, $\{f(x;\alpha,\beta) : \alpha \in (1,\infty), \beta \in (0,\infty)\}$ is an exponential family.

2d. We have $\frac{\partial w_1(\theta)}{\partial\alpha} = 1$ and $\frac{\partial w_2(\theta)}{\partial\alpha} = 0$, so that

$$\begin{aligned}
E[\log X] &= E\left[\frac{\partial w_1(\theta)}{\partial\alpha}\log X + \frac{\partial w_2(\theta)}{\partial\alpha}\log(1-X)\right] \\
&= -\frac{\partial \log c(\theta)}{\partial\alpha}
\end{aligned}$$

$$
\begin{aligned}
&= -\frac{\partial}{\partial\alpha}\{\log\Gamma[\alpha+\beta] - \log\Gamma[\alpha] - \log\Gamma[\beta]\} \\
&= -\psi(\alpha+\beta)\frac{\partial}{\partial\alpha}\{\alpha+\beta\} + \psi(\alpha) \\
&= -\psi(\alpha+\beta) + \psi(\alpha) \\
&= -\psi(2C+1) + \psi(C+1) \\
&= \{-\psi(2C+1)+\psi(2C)\} + \{-\psi(2C)+\psi(2C-1)\} + \cdots + \{-\psi(C+2)+\psi(C+1)\} \\
&= -1/(2C) - 1/(2C-1) - \cdots - 1/(C+1) \\
&= -\sum_{j=C+1}^{2C} 1/j
\end{aligned}
$$

when $\beta = \alpha - 1 = C$, a positive integer. Finally,

$$
E[X^{-1}\log X] = 2E[\log X] - E[\log X(2 - X^{-1})] = -2\sum_{j=C+1}^{2C} 1/j - 1/C = -2\sum_{j=C}^{2C-1} 1/j.
$$

3a. Let Y equal 1 with probability one. Then $E[Y] = 1$ and $P(\sqrt{Y} \geq 1) = 1$. If we can find X such that $E[X] = 0$ and $|X| = \sqrt{Y}$, then we are done since $Var[X] = E[X^2] - (E[X])^2 = E[Y] = 1$ and $P(|X| \geq 1) = P(\sqrt{Y} \geq 1) = 1$.

The requirement that $|X| = 1$ with probability one implies that X has probability mass function $f_X(x) := p1_{\{x=1\}} + (1-p)1_{\{x=-1\}}$ for some $p \in [0,1]$. Then the requirement that $E[X] = 2p - 1 = 0$ implies that $p = 1/2$. So, if X equals 1 with probability 1/2 and equals -1 with probability 1/2, then Chebychev's Inequality as applied to X is sharp.

3b. Mimicking the approach used to solve exercise 3a, we begin by seeking nonnegative Y such that $E[Y] = 1$ and $P(\sqrt{Y} \geq c) = P(Y \geq c^2) = 1/c^2$. A reasonable guess is that Y should have probability mass function $f_Y(y) := (1/c^2)1_{\{y=a\}} + (1-1/c^2)1_{\{y=0\}}$ for some number $a \geq c^2$. This yields $E[Y] = a/c^2$, which equals 1 when $a = c^2$.

Now we look for X such that $E[X] = 0$ and $|X| = \sqrt{Y}$. We see readily that X should have probability mass function $f_X(x) := 1/(2c^2)1_{\{x=-c\}} + (1-1/c^2)1_{\{x=0\}} + 1/(2c^2)1_{\{x=c\}}$. So, if X equals c with probability $1/(2c^2)$, equals 0 with probability $1 - 1/c^2$, and equals $-c$ with probability $1/(2c^2)$, then Chebychev's Inequality as applied to X is sharp.

3c. Put $Y := |X|$. Then $Y^2 = X^2$ so that $E[Y^2] = E[X^2] = Var[X] + (E[X])^2 = 1$. As such,

$$
1 = \int_0^\infty y^2 f_Y(y)\ dy = \int_0^c y^2 f_Y(y)\ dy + \int_c^\infty y^2 f_Y(y)\ dy
$$

for any $c \in (0,\infty)$. Since $\lim_{c\to\infty}\int_0^c y^2 f_Y(y)\ dy = \int_0^\infty y^2 f_Y(y)\ dy = 1$, we see that $\lim_{c\to\infty}\int_c^\infty y^2 f_Y(y)\ dy = 0$. On the other hand, since $y^2 \geq c^2$ when $y \geq c$, we have $\int_c^\infty y^2 f_Y(y)\ dy \geq \int_c^\infty c^2 f_Y(y)\ dy \geq 0$. Hence,

$$
0 = \lim_{c\to\infty}\int_c^\infty y^2 f_Y(y)\ dy \geq \lim_{c\to\infty}\int_c^\infty c^2 f_Y(y)\ dy \geq 0,
$$

which yields $\lim_{c\to\infty}\int_c^\infty c^2 f_Y(y)\ dy = 0$. The proof is completed by noting that

$$
c^2 P(|X| \geq c) = c^2 P(Y \geq c) = \int_c^\infty c^2 f_Y(y)\ dy.
$$

This shows that, asymptotically as $c \to \infty$, Chebychev's Inequality becomes less and less useful in that the probability being bounded becomes negligible compared to the bound.