# Stat 345/445: Theoretical Statistics I: Homework 5 Solutions

## Textbook Exercises

**2.26** Let $f(x)$ be a pdf and let $a$ be a number such that, for all $\epsilon > 0$, $f(a+\epsilon) = f(a-\epsilon)$. Such a pdf is said to be *symmetric* about the point $a$.

(b) (445: 1 pt.) Show that if $X \sim f(x)$, symmetric, then the median of $X$ is the number $a$.

$$\begin{aligned}
\int_a^\infty f(x)dx &= \int_0^\infty f(a+\epsilon)d\epsilon && \text{By change of variable, } \epsilon = x-a \\
&= \int_0^\infty f(a-\epsilon)d\epsilon && f(a+\epsilon) = f(a-\epsilon) \text{ for all } \epsilon > 0 \\
&= \int_{-\infty}^a f(x)dx && \text{By change of variable, } x = a-\epsilon
\end{aligned}$$

Since

$$\int_{-\infty}^a f(x)dx + \int_a^\infty f(x)dx = \int_{-\infty}^\infty f(x)dx = 1,$$

it must be that

$$\int_{-\infty}^a f(x)dx = \int_a^\infty f(x)dx = \frac{1}{2}.$$

Therefore, $a$ is a median.

**3.28** Show that each of the following families is an exponential family.

(b) (345 & 445: 2 pts.) gamma family with either parameter $\alpha$ or $\beta$ known or both unknown

$\alpha$ known,

$$f(x|\beta) = \frac{1}{\Gamma(a)\beta^a} x^{\alpha-1} e^{-\frac{x}{\beta}},$$

$$h(x) = \frac{x^{\alpha-1}}{\Gamma(\alpha)}, \quad x > 0, \quad c(\beta) = \frac{1}{\beta^\alpha}, \quad w_1(\beta) = \frac{1}{\beta}, \quad t_1(x) = -x.$$

$\beta$ known,

$$f(x|\alpha) = e^{-x/\beta} \frac{1}{\Gamma(\alpha)\beta^\alpha} \exp((\alpha-1)\log x),$$

$$h(x) = e^{-x/\beta}, \quad x > 0, \quad c(\alpha) = \frac{1}{\Gamma(\alpha)\beta^\alpha}, \quad w_1(\alpha) = \alpha - 1, \quad t_1(x) = \log x.$$

$\alpha, \beta$ unknown,

$$f(x|\alpha,\beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} \exp((\alpha-1)\log x - \frac{x}{\beta}),$$

$$h(x) = I_{\{x>0\}}(x), \quad c(\alpha,\beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha}, \quad w_1(\alpha) = \alpha - 1, \quad t_1(x) = \log x, \quad w_2(\alpha,\beta) = -1/\beta, \quad t_2(x) = x$$

(d) (345 & 445: 1 pt.) Poisson family

$$f(x) = \frac{1}{x!} I_{\{0,1,2,\cdots\}}(x)\theta^x e^{-\theta} = \frac{1}{x!} I_{\{0,1,2,\cdots\}}(x) \exp{(x\log\theta)} e^{-\theta}$$

$$h(x) = \frac{1}{x!} I_{\{0,1,2,\cdots\}}(x), \quad c(\theta) = e^{-\theta}, \quad w_1(\theta) = \log(\theta), \quad t_1(x) = x$$

(e) (345: 1 pt.) Negative binomial family with $r$ known, $0 < p < 1$

$$h(x) = \binom{x-1}{r-1} I_{\{r,r+1,\cdots\}}(x), \quad c(p) = (\frac{p}{1-p})^r, \quad w_1(p) = \log(1-p), \quad t_1(x) = x$$

**3.37** (345: 2 pts & 445: 1 pt.) Show that if $f(x)$ is a pdf, symmetric about 0, then $\mu$ is the median of the location-scale pdf $(1/\sigma)f((x-\mu)/\sigma)$, $-\infty < x < \infty$.

The pdf $(\frac{1}{\sigma})f(\frac{(x-\mu)}{\sigma})$ is symmetric about $\mu$ because, for any $\epsilon > 0$,

$$\frac{1}{\sigma} f\Big(\frac{(\mu+\epsilon)-\mu}{\sigma}\Big) = \frac{1}{\sigma} f(\frac{\epsilon}{\sigma}) = \frac{1}{\sigma} f(-\frac{\epsilon}{\sigma}) = \frac{1}{\sigma} f\Big(\frac{(\mu-\epsilon)-\mu}{\sigma}\Big)$$

Thus by Excersie 2.26b, $\mu$ is the median.

**3.39** (345 & 445: 1 pt.) Consider the Cauchy family defined in Section 3.3. This family can be extended to a location-scale family yielding pdfs of the form

$$f(x|\mu,\sigma) = \frac{1}{\sigma\pi(1+(\frac{x-\mu}{\sigma})^2)}, \quad -\infty < x < \infty.$$

The mean and variance do not exist for the Cauchy distributions. So the parameters $\mu$ and $\sigma^2$ are not the mean and variance. But they do have important meaning. Show that if $X$ is a random variable with a Cauchy distribution with parameters $\mu$ and $\sigma$, then:

(a) $\mu$ is the median of the distribution of $X$, that is $P(X \geq \mu) = P(X \leq \mu) = \frac{1}{2}$.

The pdf is symmetric about 0, so 0 must be the median. Verifying this, write

$$P(Z \geq 0) = \int_0^\infty \frac{1}{\pi}\frac{1}{1+z^2} dz = \frac{1}{\pi}\tan^{-1}(z)\Big|_0^\infty = \frac{1}{\pi}(\frac{\pi}{2} - 0) = \frac{1}{2}.$$

**3.42** (445: 2 pts.) Refer to Exercise 3.41 for the definition of a stochastically increasing family.

(a) Show that a location family is stochastically increasing in its location parameter.

Let $\theta_1 > \theta_2$. Let $X_1 \sim f(x-\theta_1)$ and $X_2 \sim f(x-\theta_2)$. Let $F(z)$ be the cdf corresponding to $f(z)$ and let $Z \sim f(x)$. Then

$$\begin{aligned} F(x|\theta_1) &= P(X_1 \leq x) = P(Z+\theta_1 \leq x) = P(Z \leq x-\theta_1) = F(x-\theta_1) \\ &\leq F(x-\theta_2) = P(Z \leq x-\theta_2) = P(Z+\theta_2 \leq x) = P(X_2 \leq x) = F(x|\theta_2). \end{aligned}$$

This inequality is because $x-\theta_2 > x-\theta_1$, and $F$ is nondecreasing. To get strict inequality for some $x$, let $(a,b]$ be an interval of length $\theta_1-\theta_2$ with $P(a < Z \leq b) = F(b)-F(a) > 0$. Let $x = a+\theta_1$. Then

$$\begin{aligned} F(x|\theta_1) &= F(x-\theta_1) = F(a+\theta_1-\theta_1) = F(a) \\ &< F(b) = F(a+\theta_1-\theta_2) = F(x-\theta_2) = F(x|\theta_2). \end{aligned}$$

**3.46** (345: 3pts & 445: 2 pts.) Calculate $P(|X - \mu x| \geq k\sigma_x)$ for $X \sim$ uniform(0,1) and $X \sim$ exponential($\lambda$), and compare your answers to the bound from Chebychev's Inequality.

Special instructions for 3.46: Give the general expression for any $k$ and then calculate the probabilities and the Chebychev bound for $k = 0.5, 1, 1.5, 2, 3$

For $X \sim$ uniform$(0,1), \mu = \frac{1}{2}$ and $\sigma^2 = \frac{1}{12}$, thus

$$P(|X-\mu| \geq k\sigma) = 1 - P\Big(\frac{1}{2} - \frac{k}{\sqrt{12}} \leq X \leq \frac{1}{2} + \frac{k}{\sqrt{12}}\Big) = \begin{cases} 1 - \frac{2k}{\sqrt{12}} & k < \sqrt{3} \\ 0 & k \geq \sqrt{3} \end{cases}$$

For $X \sim$ exponential $(\lambda), \mu = \lambda$ and $\sigma^2 = \lambda^2$, thus

$$P(|X-\mu| > k\sigma) = 1 - P(\lambda - k\lambda \leq X \leq \lambda + k\lambda) = \begin{cases} 1 + e^{-(k+1)} - e^{k-1} & k \leq 1 \\ e^{-(k+1)} & k > 1 \end{cases}$$

From Example 3.6.2, Chebychev's Inequality gives the bound $P(|X - \mu| > k\sigma) \leq 1/k^2$.

Comparison of probabilities

| $k$ | $u(0,1)$ exact | $\exp(\lambda)$ exact | Chebychev |
|---|---|---|---|
| 0.1 | 0.942 | 0.926 | 100 |
| 0.5 | 0.711 | 0.617 | 4 |
| 1 | 0.423 | 0.135 | 1 |
| 1.5 | 0.134 | 0.0821 | 0.44 |
| $\sqrt{3}$ | 0 | 0.0651 | 0.33 |
| 2 | 0 | 0.0498 | 0.25 |
| 3 | 0 | 0.0183 | 0.111 |
| 4 | 0 | 0.00674 | 0.0625 |
| 10 | 0 | 0.0000167 | 0.01 |