# Stat 5102 (Geyer) Spring 2009
# Homework Assignment 4
# Due Wednesday, February 18, 2009

Solve each problem. Explain your reasoning. No credit for answers with no explanation. If the problem is a proof, then you need words as well as formulas. Explain why your formulas follow one from another.

**4-1.** Calculate the ARE of the sample mean $\overline{X}_n$ versus the sample median $\widetilde{X}_n$ as an estimator of the center of symmetry for

(a) The Laplace location-scale family having density given in the brand name distributions handout.

(b) The $t(\nu)$ location-scale family, with densities given by

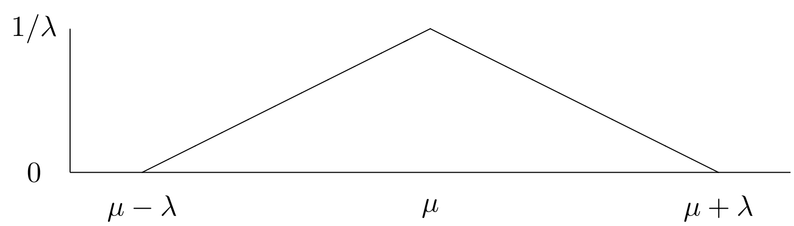
$$f_{\nu,\mu,\sigma}(x) = \frac{1}{\sigma} f_\nu\left(\frac{x - \mu}{\sigma}\right)$$

where $f_\nu$ is the $t(\nu)$ density given in the brand name distributions handout. (Be careful to say things that make sense even considering that the $t(\nu)$ distribution does not have moments of all orders. Also $\sigma$ is *not* the standard deviation.)

(c) The family of distributions called $\text{Tri}(\mu, \lambda)$ (for triangle) with densities

$$f_{\mu,\lambda}(x) = \frac{1}{\lambda}\left(1 - \frac{|x - \mu|}{\lambda}\right), \qquad |x - \mu| < \lambda$$

shown below




The parameter $\mu$ can be any real number, $\lambda$ must be positive.

**4-2.** Let $X_1$, $X_2$, $\ldots$, $X_n$ be an IID sample having the $\mathcal{N}(\mu, \sigma^2)$ distribution, where $\mu$ and $\sigma^2$ are unknown parameters, and let $S_n^2$ denote the sample variance (defined as usual with $n - 1$ in the denominator). Suppose $n = 5$ and $S_n^2 = 53.3$. Give an exact (not asymptotic) 95% confidence interval for $\sigma^2$.

**4-3.** In an experimental weight loss program five subjects were weighed before and after the 15 week treatment. The weights in pounds were as follows

| | Subject | | | | |
|---|---|---|---|---|---|
| | A | B | C | D | E |
| Before | 225 | 216 | 215 | 225 | 186 |
| After | 193 | 206 | 171 | 223 | 156 |

If you want to use R on this problem, the data are in the file

```
http://www.stat.umn.edu/geyer/5102/data/prob4-3.txt
```

(a) Calculate an exact 95% confidence interval for the expected weight loss under the program.

(b) Describe the assumptions required to make this a valid confidence interval.

**4-4.** Suppose $X_1, \ldots, X_n$ are IID having the $\text{Poi}(\mu)$ distribution and we want to make an asymptotic confidence interval for $\mu$ with coverage probability $1 - \alpha$.

(a) Give a formula for the confidence interval derived from the obvious method of moments estimator and the plug-in principle.

(b) Give a formula for the confidence interval derived from the asymptotically pivotal quantity

$$\frac{\overline{X} - \mu}{\sqrt{\mu/n}} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1).$$

(c) Give a formula for the confidence interval derived using the variance stabilizing transformation $\mu \mapsto \sqrt{\mu}$ for the Poisson distribution that was derived in 5101 homework problem 12-5.

**4-5.** Suppose we have two independent samples of size $m$ and $n$ from two different populations. We are interested in the fractions $p$ and $q$ of the populations that have a certain property (note: we are not using the $q = 1 - p$ convention here, $p$ is the proportion of the first population having the property, and $q$ is the proportion of the second population). We estimate these proportions by the sample proportions $\hat{p}_m$ and $\hat{q}_n$ which are the fractions of the first and second samples having the property. Show that

$$\hat{p}_m - \hat{q}_n \pm z_{\alpha/2} \sqrt{\frac{\hat{p}_m(1 - \hat{p}_m)}{m} + \frac{\hat{q}_n(1 - \hat{q}_n)}{n}} \tag{1}$$

is an asymptotic $100(1 - \alpha)\%$ confidence interval for $p - q$, where $z_{\alpha/2}$ is the $1 - \alpha/2$ quantile of the standard normal distribution.

**4-6.** A physics lab is divided into 20 teams. Each team performs a measurement of the speed of light. Ten teams use one method and the other ten use another method. The average and standard deviation for the teams using each method was given in the following table (units are meters per second times $10^8$).

| | mean | standard deviation |
|---|---|---|
| Method 1 | 3.00013 | 0.00395 |
| Method 2 | 2.99019 | 0.00853 |

If you want to use R on this problem, the data are in the file

```
http://www.stat.umn.edu/geyer/5102/data/prob4-6.txt
```

(a) Assuming that the measurements within each group of ten teams are independent and identically distributed around some unknown mean value (the speed of light as measured by that method), calculate a 95% confidence interval for the difference in the mean values for the two methods using Welch's approximation.

(b) Redo part (a) using the "pooled variance" $t$ confidence interval that assumes both measurement methods have the same variance.

**4-7.** Suppose a sample of size $n = 100$ is assumed to be IID having the $\text{Gam}(\alpha, \lambda)$ distribution and the method of moments estimators of the parameters are $\hat{\alpha}_n = 5.23$ and $\hat{\lambda}_n = 21.3$. Find an asymptotic 95% confidence interval for $\alpha$.

**4-8.** Suppose $V_{X,m}$ and $V_{Y,n}$ are sample variances and $M_{4,X,m}$ and $M_{4,Y,n}$ are the sample fourth central moments of independent samples from two populations having variances $\sigma_X^2$ and $\sigma_Y^2$, respectively. Find an asymptotic confidence interval for $\sigma_X^2 - \sigma_Y^2$.

**4-9.** Show that the "exact" confidence interval for the variance based on the chi-square distribution is asymptotically nonparametric distribution free within the class of all distributions having fourth moments and satisfying $\mu_4 = 3\sigma^4$. That is, show that the assumption

$$\frac{nV_n}{\sigma_2} \sim \text{chi}^2(n-1)$$

implies a certain asymptotic limit for $V_n$ and that this limit matches the correct asymptotic limit given on slide 100 of deck 1 if only if $\mu_4 = 3\sigma^4$.