Stat 5102 (Geyer) Spring 2016
Homework Assignment 1
Due Wednesday, September 14, 2016

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.

1-1. For the following data

2.2	0.8	2.0	3.6	3.6	1.8	2.2	1.0	1.2	2.4
1.4	3.2	1.0	1.8	1.6	0.4	0.2	2.0	2.2	2.4

(a) Find the mean of the empirical distribution.

(b) Find the variance of the empirical distribution.

(c) Find the standard deviation of the empirical distribution.

(d) Find the median of the empirical distribution.

(e) Find the lower and upper quartiles (the 0.25 and 0.75 quantiles) of the empirical distribution.

(f) Plot the empirical distribution function.

(g) Find $\Pr(X < 2.5)$ under the empirical distribution.

1-2. Suppose n is a positive integer, p is a real number such that $0 \le p \le 1$ and np is an integer. Suppose the data consist of $n(1-p)$ zeros and np ones. Show that

$$
\begin{aligned}
E_n(X) &= p \\
\operatorname{var}_n(X) &= p(1-p)
\end{aligned}
$$

Hint: no calculation necessary if you apply the theory you know from 5101.

1-3. The *median absolute deviation from the median* (MAD) of a random variable X with unique median m is the median of the random variable $Y = |X - m|$. The MAD of the values $x_1, \ldots, x_n$ is the median of the values $|x_i - \tilde{x}_n|$, where $\tilde{x}_n$ is the empirical median defined on Slide 20, Deck 1. The *interquartile range* of a random variable X with unique quartiles is the difference upper quartile minus lower quartile.

(a) Show that for a symmetric continuous random variable with strictly positive PDF the MAD is half the interquartile range. (The point of requiring a strictly positive PDF is that this makes all the quantiles unique and distinct.

(b) Calculate the MAD for the standard normal distribution.

(c) Calculate the MAD for the standard Cauchy distribution.

(d) Calculate the MAD for the data in Problem 1-1. (Warning: there is a `mad` function in R but it only calculates the MAD if the optional argument `constant = 1` is supplied; by default it does something else.)

1-4. Show that if X_1, ..., X_n are IID $\text{Gam}(\alpha, \lambda)$, then

$$\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$$

has the $\text{Gam}(n\alpha, n\lambda)$ distribution. Hint: use the addition rule for the gamma distribution and the change-of-variable formula for the change-of-variable $x \mapsto x/n$.

1-5. Show that if X has the $t(\nu)$ distribution, then X^2 has the $F(1, \nu)$ distribution.

1-6. Show that if X has the $F(\mu, \nu)$ distribution and $\nu > 2$, then

$$E(X) = \frac{\nu}{\nu - 2}$$

1-7. Show that if X has the $t(\nu)$ distribution and $\nu > 2$, then

$$\text{var}(X) = \frac{\nu}{\nu - 2}$$