ECON 5350 Solutions to Problem Set #2

1. The following sample is drawn from a normal distribution with mean μ and standard deviation σ:

$$X = (1.3, 2.1, 0.4, 1.3, 0.5, 0.2, 1.8, 2.5, 1.9, 3.2)\,.$$

Compute the mean, median, variance, and standard deviation of the sample. Do not use Gauss for this problem (except maybe to check your answer).

Answer.

(a) Mean = 1.52.

(b) Median = 1.55.

(c) Variance = 0.942.

(d) Standard deviation = 0.970.

2. A common method of simulating random draws from the standard normal distribution is to compute the sum of 12 draws from the $uniform[0,1]$ distribution and subtract 6. Can you justify this procedure?

Answer. Let $X_1, ..., X_{12}$ be a random sample with $X_i \sim Uniform[0,1]$. Then we know $E(X_i) = \mu = 0.5$ and $Var(X_i) = \sigma^2 = 1/12$. Let

$$Z_i = \sqrt{12}(\bar{X}_{12} - \mu)/\sigma = \sum\nolimits_{i=1}^{12} X_i - 6.$$

Therefore, by the central limit theorem, $Z \overset{asy}{\sim} N(0,1)$.

3. The random variable X has a continuous distribution $f(x)$ and cumulative distribution function $F(x)$. What is the probability distribution of the sample maximum? [Hint: In a random sample on n observations, $\{X_1, X_2, \ldots, X_n\}$, if Z is the maximum, then every observation in the sample is less than or equal to Z. Use the cdf.]

Answer. Let Y_n be the sample maximum of $X_1, ..., X_n$. The probability that every sample observation is less than or equal to Y_n is

$$\Pr(X_1 \leq y_n, X_2 \leq y_n, ..., X_n \leq y_n) = F(y_n)^n.$$

The probability density function is then

$$\frac{dF(y_n)^n}{dy_n} = nF(y_n)^{n-1} f(y_n).$$

4. In random sampling from the exponential distribution

$$f(x) = (1/\theta)exp[-x/\theta],\ x, \theta > 0,$$

find the maximum likelihood estimator of θ and obtain the asymptotic distribution of this estimator.

Answer. The likelihood for the random sample $X_1, ..., X_n$ from $X_i \sim Exponential(\theta)$ is

$$L(\theta) = \prod_{i=1} \frac{1}{\theta} \exp(-x_i/\theta) = \left(\frac{1}{\theta}\right)^n \exp(-\frac{1}{\theta} \sum_{i=1}^n x_i).$$

Take natural logs of $L(\theta)$:

$$\ln L(\theta) = -n \ln(\theta) - \frac{1}{\theta} \sum_{i=1}^n x_i.$$

The first and second derivatives are

$$\begin{aligned} \frac{d \ln L}{d\theta} &= -\frac{n}{\theta} + \frac{1}{\theta^2} \sum_{i=1}^n x_i \\ \frac{d^2 \ln L}{d\theta^2} &= \frac{n}{\theta^2} - 2\frac{1}{\theta^3} \sum_{i=1}^n x_i. \end{aligned}$$

Therefore, the maximum likelihood estimator is $\hat{\theta}_{ML} = \bar{X}$. The asymptotic variance of $\hat{\theta}_{ML}$ is

$$I^{-1}(\theta) = -E\left(\frac{d^2 \ln L}{d\theta^2}\right)^{-1} = -E\left(\frac{n}{\theta^2} - 2\frac{n}{\theta^3}\bar{X}\right)^{-1} = -\left(\frac{n}{\theta^2} - 2\frac{n}{\theta^3}E(\bar{X})\right)^{-1} = \frac{\theta^2}{n}.$$

Therefore, we know that $\hat{\theta}_{ML} \overset{asy}{\sim} N(\theta, \theta^2/n)$.

GAUSS PROBLEMS

5. Let $\{X_1, X_2, \ldots, X_{20}\}$ be a random sample from a $uniform[0, 1]$ distribution.

   (a) Generate a histogram to estimate the sampling distribution for Y_1, the first order statistic, by simulating 500 random samples (each of size 20).

   (b) Calculate the actual sampling distribution for Y_1 ($n = 20$) and create an xy graph of the pdf. Comment on the similarities and differences between the actual and estimated sampling distributions.

Answer.

(a) See attached gauss code.

(b) The probability that every sample observation is greater than or equal to Y_1 is

$$\Pr(X_1 \geq y_1, X_2 \geq y_1, ..., X_n \geq y_1) = (1 - F(y_1))^{20}.$$

The sampling density is then

$$20\,(1-F(y_1))^{19}\,f(y_1)=g_1(y_1).$$

Substituting in the cdf and pdf for the $Uniform[0,1]$ distribution results in

$$g_1(y_1)=20(1-y_1)^{19},\ 0\leq y_1\leq 1$$

and zero otherwise.

6. Maximum Likelihood and Hypothesis Testing.

   (a) Find the MLE of λ assuming that the time-series data in hw2q8.txt are generated from a Poisson(λ) distribution ($n=11$). The data are the number of homicides in New York City from 1985-1995 as reported by the Bureau of Justice Statistics.

   (b) What is the estimated asymptotic variance of the ML estimate of λ? Use the result from question 5 of problem set #1 to calculate the exact sample variance. Comment on the performance of the asymptotic variance relative to the exact variance. Also, comment on the difference.

Answer.

   (a) See attached gauss code.

   (b) We know from question 5 of problem set #1 that the sum of Poisson random variables has a Poisson distribution. Consequently,

$$\hat{\lambda}_{ML}=\bar{X}\sim Poisson(\mu=\lambda,\sigma^2=\lambda/n).$$

   To summarize, $\hat{\lambda}_{ML}$ has an asymptotic variance equivalent to its finite-sample variance. Its asymptotic distribution is normal, but it has a Poisson distribution in finite samples.

```
@ *********************** @
@ ECON 5350 PROBLEM SET #2 @
@ *********************** @

@ ********** @
@ PROBLEM #5 @
@ ********** @

@ Part (a) @
loops = 500;
nobs = 20;
y1 = zeros(loops,1);
counter = 1;

do while counter <= loops;
        x = rndu(nobs,1);
        minimum = minc(x);
        y1[counter,1] = minimum;
        counter = counter +1;
endo;

library pgraph;
graphset;
title("Simulated Sampling Distribution of Y1 from a Uniform(0,1) pdf");
{b,m,f} = hist(y1,20);

@ Part (b) @
steps = 100;
gy = zeros(steps,1);
y1axis = seqa(0.01,0.01,steps);
counter1 = 1;

do while counter1 <= steps;
        gy[counter1,1] = nobs*(1 - y1axis[counter1,1])^(nobs-1);
        counter1 = counter1 + 1;
endo;

graphset;
title("Actual Sampling Distribution of Y1 from a Uniform(0,1) pdf");
xy(y1axis,gy);

@ ********** @
@ PROBLEM #6 @
@ ********** @

@ Loading in the Data @
load path = c:\gauss35\classes\econ5340\data\;
load x[12,1] = hw2q8.txt;
data = x[2:12,1];
nobs = rows(data);

@ Part (a) @
URlambda = meanc(data);
print "Maximum Likelihood Estimate of Lambda = " URlambda;

@ Part (b) @
asyvar = URlambda/nobs;
exactvar = asyvar;
print "Estimated Asymptotic Variance of the ML Estimate = " asyvar;
print "Estimated Exact Variance of the ML Estimate = " exactvar;

Maximum Likelihood Estimate of Lambda =        1774.2727
Estimated Asymptotic Variance of the ML Estimate =        161.29752
```

```
Estimated Exact Variance of the ML Estimate =        161.29752
```