CALIFORNIA STATE UNIVERSITY, EAST BAY
STATISTICS DEPARTMENT

Statistics 6501 Mathematical Statistics
Winter 2011

Take-home Midterm

Instructions: This is the take-home part of the test. This is a test. You are to work on this test alone and you are not to talk with others in the class. This take-home part of the test will be due next week on Monday.

Simulation in R

1. Simulate from the Cauchy distribution 10,000 times, three different ways.

   (a) Generate 10,000 random values from the $Unif(-\frac{\pi}{2}, \frac{\pi}{2})$.
   Run the following R code.

   ```
   n = 100000
   u = runif(n, min=-pi/2, max=pi/2)
   x = tan(u)
   index = 1:n
   x.cummean = cumsum(x)/index
   plot(x.cummean, type = "l")
   ```

   (b) Generate two vectors of 10,000 random values from the $N(0, 1)$.
   Run the following R code.

   ```
   n = 10000
   x = rnorm(n)
   y = rnorm(n)
   w = x/y
   index = 1:n
   w.cummean = cumsum(w)/index
   plot(w.cummean, type = "l")
   ```

   (c) Generate two vectors x1 and x2 of 10,000 random values from the $N(0, 1)$.
   Run the following R code.

   ```
   n = 10000
   x1 = rnorm(n)
   x2 = rnorm(n)
   y1 = x1 + x2
   y2 = x1 - x2
   plot(y1,y2)   # Does the plot look uncorrelated?
   x.bar = (x1+x2)/2
   s2 = (x1-x2)**2/2
   plot(x.bar, s2)  \# Does the plot look uncorrelated?
   t.stat = sqrt(2)*x.bar/sqrt(s2)
   index = 1:n
   t.stat.cummean = cumsum(t.stat)/index
   plot(t.stat.cummean, type = "l")
   ```

2. Simulate from the general bivariate normal distribution and transform to independence.

   Start by examining the handout BVNsim.R to answer the following questions.

   (a) Make a plot of the $BVN(\mu_1 = 10, mu_2 = 25, \sigma_1^2 = 2^2 = 4, \sigma_2^2 = 3^2 = 9, \rho = -0.4)$.

   (b) Simulate two vectors of $Unif(0,1)$ random values of length 2,000. Make histograms of each vector of random values and make a scatterplot, one vector on the x-axis and the other on the y-axis.

   (c) Transform the uniform random values to independent standard normal random values using the Box-Muller method. Make histograms of each vector of random values and make a scatterplot, one vector on the x-axis and the other on the y-axis.

   (d) Transform the $BVN(0,0,1,1,0)$ to $BVN(0,0,\sigma_1^2,\sigma_2^2,\rho)$. Make histograms of each vector of random values and make a scatterplot, one vector on the x-axis and the other on the y-axis.

   (e) Transform the $BVN(0,0,\sigma_1^2,\sigma_2^2,\rho)$ to $BVN(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$. Make histograms of each vector of random values and make a scatterplot, one vector on the x-axis and the other on the y-axis.

   (f) Determine the angle of rotation θ to transform the BVN to independence. Rotate $BVN(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$ to $BVN(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,0)$. Make histograms of each vector of random values and make a scatterplot, one vector on the x-axis and the other on the y-axis.

   (g) Use the R function ipairs( ) in the IDPmisc library to make better scatterplots.

   (h) Use the R function hist2d( ) int he gplots library to make 2 dimensional histograms.

3. Simulate from the bivariate p.d.f. of the minimum and maximum sampling from the $N(0,1)$. Make a scatterplot of the mininums versus the maximums.

   Simulate from the bivariate p.d.f. of the minimum and maximum sampling from the $Unif(0,1)$. Make a scatterplot of the minimums versus the maximums.

   Does increasing the sample size change the correlation between the minimum and the maximum?

```
n = 100  # try 5 and 100 also
B = 10000
x.min = numeric(B)
x.max = numeric(B)
for (i in 1:B){
    x = rnorm(n)    # x = runif(n)
    x.min[i] = min(x)
    x.max[i] = max(x)
}
plot(x.min,x.max)
cor(x.min,x.max)
```