Risk and Portfolio Management with Econometrics, Courant Institute, Spring 2009
http://www.math.nyu.edu/faculty/goodman/teaching/RPME09/index.html
Always check the class bboard on the blackboard site from home.nyu.edu (click on academics, then on the course name) before doing any work on the assignment.

Assignment 2, due February 5

Corrections: (none yet)

1. Suppose $Z \in R^n$ is a multivariate standard normal. That means that the joint density is given by $f(z) = \frac{1}{(2\pi)^{n/2}} \exp\left(\frac{-1}{2}\sum_{k=1}^n z_k^2\right)$. In this problem, $\|z\|$ always represents the Euclidean norm $\|z\|^2 = \sum_{k=1}^n z_k^2$.

   (a) Show that $\|Z\|^2 = \sum_{k=1}^n Z_k^2$ has a χ_n^2 distribution. (This is for emphasis, since it's almost the definition of χ_n^2.)

   (b) Let $V \subset R^n$ be a linear subspace of dimension m, and let P_V be the orthogonal projection onto V. If $z \in R^n$ and $w = P_V z \in V$, show that w solves the minimization problem

   $$\min_{w \in V} \|w - z\|^2 \ .$$

   (Again for emphasis, since this is closely related to the definition of orthogonal projection.)

   (c) Let $W = P_V Z$, with Z as above. Show that W is a standard normal random variable in V in the same sense in which Z is a standard normal in R^n (covariance $= I_{m\times m}$). Use this to show that $\|W\|^2 = \sum_{k=1}^n W_k^2$ has a χ_m^2 distribution. It may help to think of an orthonormal basis of m vectors in V.

   (d) Let $x = (x_1, \ldots, x_n) \in R^n$ be given fixed vector (not random) and let V be the set of $z \in R^n$ orthogonal both to both x and $\mathbf{1}$. Let $W = P_V Z$ as above. Show that we can write $W = Z - a\mathbf{1} - bx$, where a and b depend on Z. Find formulas for a and b in terms of x and Z, Hint: minimize over a and b the expression

   $$\|W\|^2 \ = \ \sum_{k=1}^n (Z_k - a - bx_k)^2$$

   by setting the derivatives with respect to a and b to zero. This is one way to look at standard least squares regression.

   (e) Show that a and b are Gaussian random variables. Find their means, variances, and covariance. Recall that Z is random and x is not random. The answers may depend on x, but not on Z.

   (f) Show that $\sum_{k=1}^n (Z_k - a - bx_k)^2 \sim \chi_{n-2}^2$, using the optimal a and b from part (d).

2. Suppose X and Y are jointly normal with mean zero, variances σ_x^2 and σ_Y^2, and covariance σ_{XY}. (X and Y are each one dimensional.) Write the joint probability density $f(x,y)$. Use this to show that

$$E\left[X^2Y^2\right] \;=\; \int\!\!\int x^2y^2 f(x,y)\,dxdy \;=\; \sigma_x^4 \,+\, \sigma_Y^4 \,+\, \sigma_{xy}^2\,. \qquad (1)$$

This is an example of a general formula called *Wick's theorem*. The purpose of this exercise is to give you practice manipulating Gaussian probability densities.

3. Suppose $X \in R^n$ and $Y \in R^m$ are jointly normal with $(n+m)\times(n+m)$ covariance matrix

$$\sigma \;=\; \begin{pmatrix} \sigma_{XX} & \sigma_{XY} \\ \sigma_{YX} & \sigma_{YY} \end{pmatrix} \;.$$

This is a 2×2 *block matrix*, with each entry representing a matrix. The top left entry is the $n\times n$ covariance matrix of X, The bottom right entry is the $m\times m$ covariance matrix of $Y$with σ_{XX}. The top right entry contains covariances between X and Y components:

$$\sigma_{XY,jk} = \mathrm{cov}\left(X_j, Y_k\right) \;.$$

This matrix has one row for each component of X and one column for each component of Y, which makes it an $n\times m$ matrix. The lower left matrix is the transpose of this. You should check that this σ is the covariance matrix of the $n+m$ component vector (X,Y).

   (a) What is the marginal distribution of Y if we do not observe X? It will be easier to describe it as a certain joint normal distribution than to write the probability density function.

   (b) What is the conditional distribution of Y given an observation of X? This may be written as a density $f(y \mid X)$, but it again is easier to describe it as a multivariate normal whose parameters depend on X in some way.

   (c) It the context of Bayesian statistics, suppose that $X \sim \mathcal{N}(\mu,1)$, were we have a prior on μ that is $\mu \sim \mathcal{N}(\mu_0,\sigma_\mu^2)$, and a conditional distribution that given μ, $X \sim \mathcal{N}(\mu,\sigma_X^2)$. In the general Bayesian framework μ is called θ and the distribution of X given μ is $f(x \mid \mu)$. We are saying $f(x \mid \mu) = \mathcal{N}(\mu,\sigma_X^2)$. Here we interpret μ as a random variable rather than an unknown constant. Suppose that we first choose $\mu \sim \mathcal{N}(\mu_0,\sigma_\mu^2)$, then choose $X \sim \mathcal{N}(\mu,\sigma_X^2)$. Show that the resulting joint density of (X,μ) is normal in the sense above, with μ playing the role of Y.

   (d) Suppose instead we choose $X = (X_1,\ldots,X_n)$, with each component from the $\mathcal{N}(\mu,\sigma_X^2)$ distribution, but with the X_k independent. Show that this (X,μ) falls under the above general framework. Describe the *posterior* distribution of μ given the n *observations* $X = (X_1,\ldots,X_n)$.

(e) Show that the posterior variance of μ (the variance of $\mu \mid X$) is on the order of $1/n$.

(f) The parameter σ_μ represents our *prior* confidence of our knowledge of μ. Show that the answer to part (d) has the property that the mean of the posterior distribution converges to the sample mean as $\sigma_\mu \to \infty$. In this way, the Bayesian estimate converges to the usual one if the the case of a *flat* prior – one with no information.

(g) Show that the posterior distribution of μ converges to μ_0 as $\sigma_\mu^2 \to 0$. How do you interpret this?

4. In Matlab you use `rand` to generate independent random variables uniformly distributed in $[0, 1]$. Suppose U is such a random variable and $T = -\ln(U)/\lambda$. Show that T is an exponential random variable with *rate* parameter λ. The probability density of such a random variable is $f(t) = \lambda e^{-\lambda t}$ if $t > 0$, and $f(t) = 0$ if $t < 0$. In Matlab, generate a large number (maybe a million) of such independent exponentials, make a histogram, and show that the histogram has the right shape. Meucci's book has some hints about histograms. In particular, you need to make the bin size not so large that you don't see the distribution and not so small that the heights of the individual bars are very noisy. If Δt is bin size, the expected number of samples in a given bin is approximately $nf(t)\Delta t$, where t is the midpoint of the bin and n is the number of samples in all. Have Matlab put the histogram and the expected heights on the same graph so that you can see that the random variables are being generated correctly. Use one or two different λ values.

5. Calculate μ_T and σ_T^2, the mean and variance of the exponential random variable of Question 4. For each m, define the random variable

$$R_m = \frac{1}{\sqrt{m}} \sum_{j=1}^{m} (T_j - \mu_T) \; .$$

For large m, the distribution of R_m should (according to the Central Limit Theorem) be approximately normal with mean zero and variance σ_T^2. For various values of m, make a plot similar to that of Question 4, with the histogram and the probability distribution. Put the normal approximation on the same plot. For small m (2, 5, ..), the fit will not be very good, the histogram values will not be so well fit by the normal approximation. For larger m the fit should improve. Hand in a few plots that illustrate improving fit as m increases.

6. Suppose $V = 2U - 2$ is uniformly distributed in $[-1, 1]$ and we use it to make the two component random variable $X = V$, $Y = V^3$. Compute the 2×2 variance/covariance matrix of (X, Y). Let $R_m = \frac{1}{\sqrt{m}} \sum_{k=1}^{m} V_k$ and $S_m = \frac{1}{\sqrt{m}} \sum_{k=1}^{m} V_k^3$. Use Monte Carlo simulation to estimate $A_m = E[R_m^2 S_m^2]$ for various values of m both small and large. Show that for large values of m, the result is approximately given by the normal approximation you can calculate from Question 2. Use a large number of samples (maybe a million) for each value of m.