Assignment 4.

Given September 26, due October 10 (note: two weeks)
Last revised, October 7.

Objective: Continuous conditioning, Gaussian random variables.

1. Suppose $X \sim \mathcal{N}(\mu, \sigma^2)$. Find a formula for $E[e^X]$. Hint: write the integral and complete the square in the exponent.

2. In finance, people often use $N(x)$ for the CDF (cumulative distribution function) for the standard normal. That is, if $Z \sim \mathcal{N}(0,1)$ then $N(x) = P(Z \leq x)$. Suppose $S = e^X$ for $X \sim \mathcal{N}(\mu, \sigma^2)$. Find a formula for $E[\max(S, K)]$ in terms of the N function. Hint: as in problem 1. This calculation is part of the Black–Scholes theory of the value of a vanilla European style call option. K is the known strike price and S is the unknown stock price.

3. Suppose $X = (X_1, X_2, X_3)$ is a 3 dimensional Gaussian random variable with mean zero and covariance

$$E[XX^*] = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix} .$$

Set $Y = X_1 + X_2 - X_3$ and $Z = 2X_1 - X_2$.

a. Write a formula for the probability density of Y.

b. Write a formula for the joint probability density for (Y, Z).

c. Find a linear combination $W = aY + bZ$ that is independent of X_1.

d. Find a (modern style) formula for $E[Y \mid Z]$.

e. Interpret your answer to part d in classical style as $E[Y \mid 2X_1 - X_2 = z]$.

4. In some applications, linear filtering may be thought of as trying to find the best approximation to an unknown signal that has been corrupted by noise. Much of the theory of linear filtering rests on models using linear recurrence relations and Gaussian noise. Here is a simple example of that theory. We model the random process representing the signal as a sequence, X_n, which satisfies the recurrence relation

$$X_{n+1} = \alpha X_n + Z_n ,$$

where α is a "decay" parameter, $|\alpha| < 1$, and the Z_n are independent standard normal random variables. Assume that $X_0 = 0$.

a. For any T, show that $X = (X_1, \ldots, X_T)$ is jointly Gaussian. You need not calculate the $T \times T$ covariance marix.

The noisy observations Y_n are modeled as

$$Y_n = X_n + \beta W_n \ ,$$

where the W_n are standard normals independent of the X_m and each other, β is a noise parameter. Let $\mathcal{F}_n$ be the $\sigma-$algebra generated by the first n observations $Y_1, \ldots, Y_n$.

b. Compute $\widehat{X_n} = E[X_n \mid \mathcal{F}_n]$. This is a best (least squares) estimate of X_n from the noisy observations Y_k for $1 \leq k \leq n$. Hint: first find $\widehat{X_1}$, then find a formula for $\widehat{X}_{n+1}$ in terms of $\widehat{X_n}$ and Y_{n+1}.

Remark (for statisticians): The expected value of a gaussian is the same as the maximum likelihood estimate because the maximum of the probability density as the mean value. This might not be true for other random variables.

5. Suppose we have independent random variables $Z_k \sim \mathcal{N}(0,1)$ and Y_k with density $u(y)$ so that $u(y) = 3y^2$ if $0 \leq y \leq 1$ and $u(y) = 0$ otherwise. Consider the sums $S_n = \sum_{k=1}^n Z_k$, $T_n = \sum_{k=1}^n Z_k^2$, $V_n = \sum_{k=1}^n Y_k$, and $W_n = \sum_{k=1}^n Y_k^2$. Note that Z_k is hardly independent of Z_k^2. Use the central limit theorem to decide which is more true:

a. For large n, S_n is nearly independent of T_n.

b. For large n, V_n is nearly independent of W_n.