Midterm: SIO 221B, Data Analysis

due Friday, October 25, 2002
Open book. Open note. Do not discuss these questions. If you have questions or concerns, please contact me by e-mail.

1. An oceanographic experiment measures 3 variables, x, y, and ϕ, with measurement errors σ_x, σ_y, and σ_ϕ respectively. From these variables, you wish to compute:

$$u = x\sin(\phi) + y^2\cos(\phi). \tag{1}$$

What is σ_u, the estimated error in u? What assumptions have you used to derive σ_u?

2. Consider the following pdf:

$$F_Y(y)dy = \begin{cases} \alpha(1+y)dy & \text{for } -1 < y \leq 0 \\ \alpha dy & \text{for } 0 < y \leq 1 \\ \alpha(2-y)dy & \text{for } 1 < y \leq 2 \\ 0. & \text{otherwise} \end{cases} \tag{2}$$

a. What should α be?
b. What are the moments of the pdf F_Y (about the mean)?
c. How do you generate data with a distribution $F_Y(y)$ from a uniform distribution?

3. The quantity y is defined so that:

$$y(n) = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{3}$$

a. The variables x are uniformly distributed. What is the distribution of $y(n)$?
b. When n is large, if x has standard deviations σ_x, what is the standard deviation of y?

4. On the course web site (under homework), you will find 3 data files containing the u and v components of wind at 3 locations in the tropical Pacific.

a. Compute the covariance matrix for these winds.
b. What is the correlation between these winds?
c. What fraction of the variance in u can be explained by v? What fraction of the variance in u at one location can be explained by u at a different location? Could you use this to fill gaps in the time series?

The data are derived from raw scatterometer measurements. The QuikSCAT scatterometer measures at roughly 6 am and 6 pm local time, but it does not sample every location twice a day. Therefore the data have many gaps which are marked as -999. Be sure to avoid the -999s when you compute your solutions, but use as many observations as possible. Please describe in words the method that you use to compute the covariance.