Midterm: MAE 127 (Solutions)

Friday, April 29, 2005

1. Wind stress measures the transfer of momentum from the atmosphere to the ocean. The x-component is calculated as:

$$\tau_x = C_D(u^2 + v^2)u.$$

where C_D is a constant, and u and v are measured quantities with uncertainties δ_u and δ_v respectively.

(a) Use error propagation to estimate the uncertainty in τ_x.

Using error propagation we find:

$$\delta_\tau = C_D\sqrt{(3u^2\delta_u + v^2\delta_u)^2 + (2uv\delta_v)^2}$$

(b) What assumptions does one make by estimating error this way?

Error propagation assumes that the errors in u and v are statistically independent, that the variations in τ are essentially linear as a function of u and v. In addition, our interpretation of the errors often takes advantage of the assumption that errors in u and v are normally distributed (or Gaussian).

2. Suppose that the following 5 data points were collected, corresponding to sea level changes (in meters) in La Jolla Shores: $x = [0\ 1\ 2\ 1\ 0]$.

(a) What is the median sea level?

The median is the middle value, when sorted: 1 m.

(b) What is the mean sea level?

The mean is $\overline{x} = \sum_{i=1}^{5} x_i/5 = 0.8$ *m.*

(c) How would you compute the standard deviation?

The standard deviation is $\sigma = \sqrt{\sum_{i=1}^{5}(x_i - 0.8)^2/4}(= \sqrt{0.7} \approx 0.84)$ *m.*

(d) What is the error of the mean of x?

The error of the mean is $\sigma/\sqrt{N} = \sigma/\sqrt{5}(\approx 0.37)$ *m.*

3. For the following probability density function:

$$P(x)\,dx = \begin{cases} a\,dx & \text{for } 0 \le x < 1 \\ 2a\,dx & \text{for } 1 \le x < 2 \\ 0. & \text{otherwise} \end{cases}$$

(a) Find a.

To find a, we set the integral of the PDF between $-\infty$ *and* $+\infty$ *to 1:*

$$\int_{-\infty}^{\infty} P(x)\,dx = \int_0^1 a\,dx + \int_1^2 2a\,dx = ax|_0^1 + 2ax|_1^2 = a + 4a - 2a = 3a = 1,$$

so $a = 1/3$.

(b) Find the mean of x.

The mean of x *is determined from the integral of* $x\,P(x)\,dx$*:*

$$\langle x \rangle = \int_{-\infty}^{\infty} xP(x)\,dx = \int_0^1 ax\,dx + \int_1^2 2ax\,dx = \left.\frac{ax^2}{2}\right|_0^1 + \left.\frac{2ax^2}{2}\right|_1^2 = \frac{1}{6} + \frac{4}{3} - \frac{1}{3} = \frac{7}{6}.$$

(c) Find the median of x.

The median is determined by finding the value of x *for which the cumulative distribution function is 0.5. By inspection, we can tell that one third of the area under the PDF is between 0 and 1 and two-thirds of the area is between 1 and 2, so the median* m *must fall between 1 and 2.*

$$0.5 = \int_{-\infty}^{m} P(x)\,dx = \int_0^1 \frac{1}{3}\,dx + \int_1^m \frac{2}{3}\,dx = \left.\frac{x}{3}\right|_0^1 + \left.\frac{2x}{3}\right|_1^m = \frac{1}{3} + \frac{2m}{3} - \frac{2}{3}.$$

This implies that:

$$\frac{2m}{3} = \frac{1}{2} + \frac{1}{3} = \frac{5}{6}.$$

Thus m*, the median of* x*, is 5/4, and is not equal to the mean.*

4. For the three artificial time series (a-c) in Figure 1, estimate the correlation coefficient r. Assuming that a data pairs are collected at the times of each black circle, which of these would you guess to be statistically significant (and why)?

(a) In the first case the solid and dashed lines have equal size fluctuations of opposite sign, so the correlation coefficient must be $r = -1$*. There are 13 data points indicated by black dots, so this implies that the results are statistically significant.*

(b) In this case, the solid line is zero when the dashed line is non zero, and vice-versa, so the correlation coefficient must be $r = 0$*. This is not statistically significant.*

(c) In this case, the solid and dashed lines have different means and different amplitudes, but their fluctuations are of opposite sign, so the correlation coefficient must be $r = -1$ *as in part (a). Again, with 13 data points, we would infer this to be statistically significant.*

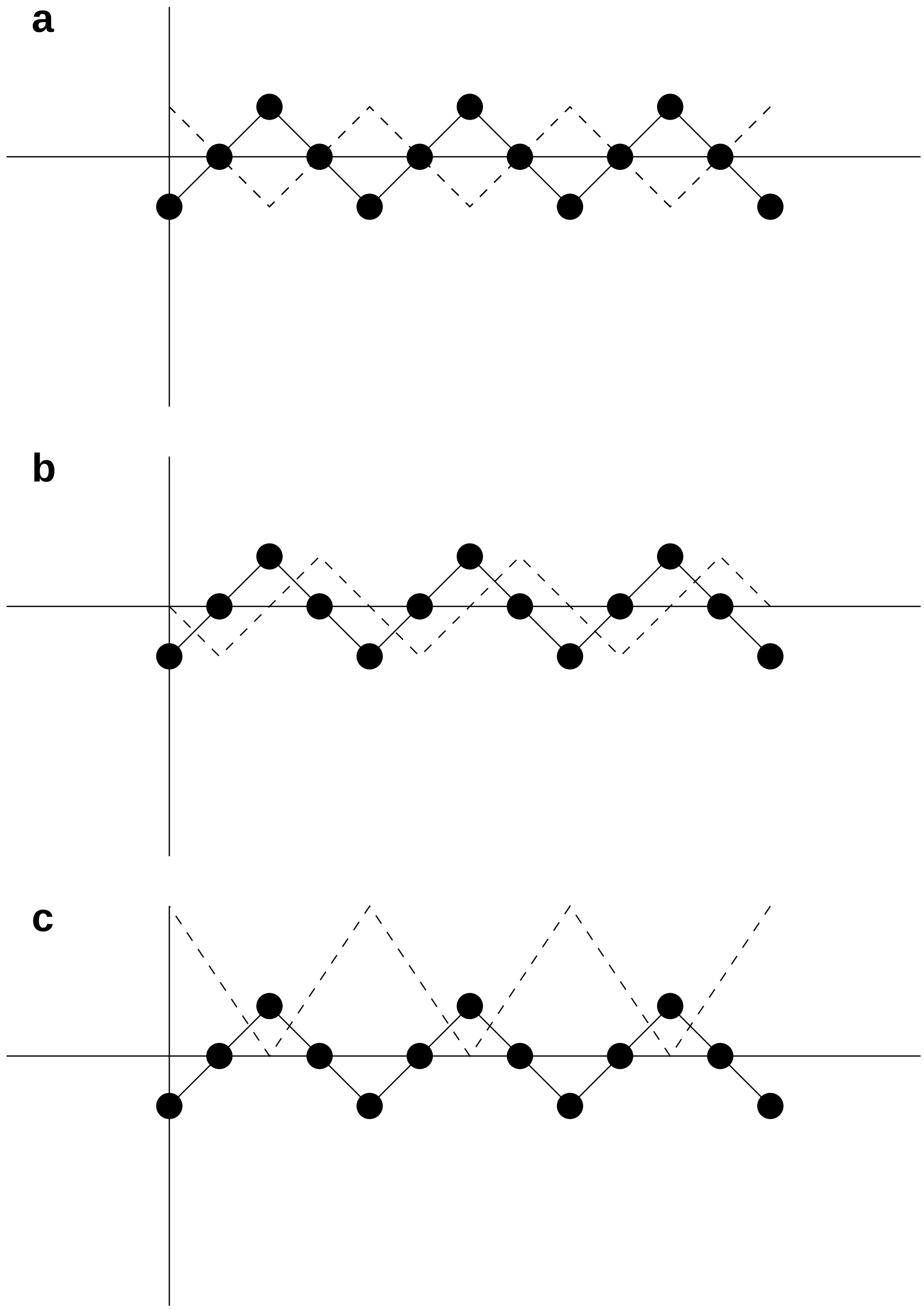

Figure 1: Artificial time series identified by solid and dashed lines. Assume that data are collected at points in time corresponding to black dots, and that each observation is statistically independent.