More Statistics

2-2-2011

Opening Discussion

- What did we talk about last class?
- Minute essay comments:
  - How do you implement these measures in code?
  - Suggested references for those who aren't up on their stats.
  - How do we generate the random numbers?
  - http://en.wikipedia.org/wiki/Australian_50_cent_coin
  - Benefit of cummulative distributions.

Correlation and Covariance- Stationary

- Correlation:
  - $$\rho_{ij} = \frac{C_{ij}}{\sqrt{\sigma_i^2 \sigma_j^2}}$$
- This allows some correlation, only the mean and the variance can't change over time.
  - $$\rho_j = \frac{C_{i,i+j}}{\sqrt{\sigma_i^2 \sigma_{i+j}^2}} = \frac{C_j}{\sigma^2} = \frac{C_j}{C_0}$$

Estimation of Means and Variance

- Given a random variable in an experiment, taking the mean or the variance of the sequence only gives an estimate of the true mean or variance.

$$\bar{X}(n) = \hat{\mu} = \frac{\sum_{i=1}^{n} X_i}{n}$$

$$S^2(n) = \hat{\sigma} = \frac{\sum_{i=1}^{n} [X_i - \bar{X}(n)]^2}{n-1}$$

- The variance in the mean, which tells us the spread in its distribution, goes as σ^2/n.

Problems

- If the values are correlated, this is still a good estimate of the mean, but the variance is off.

$$E[S^2(n)] = \sigma^2 \left[1 - 2 \frac{\sum_{j=1}^{n-1} (1 - j/n) \rho_j}{n-1} \right]$$

$$Var[\bar{X}(n)] = \sigma^2 \frac{\left[1 + 2 \sum_{j=1}^{n-1} (1 - j/n) \rho_j \right]}{n}$$

Central Limit Theorem

- For X pulled from any distribution, the distribution of the following will approach the standard normal random distribution.

$$[\bar{X}(n)-\mu]/\sqrt{\sigma^2/n} \approx [\bar{X}(n)-\mu]/\sqrt{S^2(n)/n}$$

Confidence Intervals

- We can specify an interval about the experimental mean into which the true mean should fall with a specified confidence level.

$$\bar{X}(n) \pm t_{n-1,1-\alpha/2}\sqrt{\frac{S^2(n)}{n}}$$

- The odds of falling into this interval are 1-α.
- Note how the range varies with n.

Hypothesis Test for Means

- Null hypothesis is that $\mu=\mu_0$.
- $t_n=[\bar{X}(n)-\mu_0]/\sqrt{S^2(n)/n}$
- If $t_n > t_{n-1,1-\alpha/2}$ reject hypothesis.

Strong Law of Numbers

- The mean of an experiment will approach the true mean with probability 1 as n goes to infinity.
- (Enough data points will give you an answer arbitrarily close to the real answer.)

Danger of Replacing a Distribution with Mean

- Don't do this. It leads to results that can be very misleading.
- Two distributions with similar means can lead to interesting dynamics while just the means doesn't.

Minute Essay

- Thoughts?