2001 Honors Examination in Statistics

Department of Mathematics and Statistics
Swarthmore College

Instructions: This examination consists of a total of four questions. Number the questions clearly in your work and start each question on a new page. You must make it clear how you arrived at your answer. Answers without any work may lose credit even if they are correct.

This is a closed-book three-hour examination. You may not refer to notes or textbooks.

1. The diameter of steel rods coming off an assembly line are known to be normally distributed with mean 20 centimeters and variance 4 centimeters. Based on this information, answer the following questions. You may leave your answers in terms of the Normal cumulative density function Φ.

(1a) What is the probability of observing a randomly selected steel rod with diameter greater than 23 centimeters?

(1b) Assuming rods coming off the assembly line are independent, what is the probability that 4 out of the next 6 rods will have diameter greater than 22 centimeters?

(1c) How many rods, in expectation, will I have to measure to get the first rod that has diameter greater than 24 centimeters?

In practice, it is unlikely that rods coming off an assembly line are independent, instead it is more likely that there is a positive dependence in their diameters.

(1d) What is the effect of positive dependence on the variance of the sum of the diameters of 2 successive rods, compared to if they had been independent?

(1e) Assuming that the correlation among successive rod diameters is 0.7, what is the variance of the difference in diameters between 2 rods?

2. This question has five parts as below.

(2a) Write down the probability density function for the number of failures x before the <u>first</u> success in a series of independent coin flips with constant probability. Use the notation p = probability of success and q = 1-p, the probability of failure. (This is called the Geometric distribution).

(2b) Given that the moment generating function for the Geometric distribution is

$$M(t) = pe^t/(1-qe^t)$$

compute the mean and variance of the Geometric distribution.

(2c) Write down the probability density function for the number of failures x before r successes are observed in a series of independent coin flips with constant probability. (This is called the Negative Binomial Distribution).

(2d) If $p = 0.7$, using your answer to (2c), what is the probability of observing at least one failure before 2 successes? (Note, even if you can't answer (2c), you can derive this answer in another manner.)

(2e) Other that the process described in (2c), how else does the Negative Binomial Distribution naturally arise? (Hint: The Negative Binomial Distribution is an overdispersed version of another common distribution.)

3. Consider the following Bayesian Hierarchical Model:

$$Y_{ij} \sim N(\theta_i, \sigma^2_i) \qquad i=1,\ldots,I; \qquad j=1,\ldots,n_i$$

$$\theta_i \sim N(\bar{\theta}, \tau^2)$$

where i indexes individuals, and $j=1,\ldots, n_i$ indexes observations within individuals.

(3a) What is an unbiased estimate of $\bar{\theta}$?

(3b) What is the variance of your estimator in part (3a) (you may assume for this answer that σ^2_i and τ^2 are known)?

(3c) Derive the posterior distribution of θ_i conditional on the values of Y_{ij}, σ^2_i, τ^2, and $\bar{\theta}$ (that is $p(\theta_i | Y_{ij}, \sigma^2_i, \tau^2, \bar{\theta})$).

(3d) Suppose you could obtain samples from the joint posterior distribution $p(\sigma^2_i, \tau^2, \bar{\theta} | Y_{ij})$, explain how you could use samples from this distribution, in conjunction with the posterior distribution derived in (3c), to obtain samples from $p(\theta_i | Y_{ij})$.

(3e) Explain how to obtain an unbiased estimate of τ^2 using the set of s^2_i, the sample variances within person, and s^2, the sample marginal variance of the Y_{ij}. Relate your answer to a repeated measures analysis-of-variance experimental design.

4. Let $y_1, \ldots, y_n$ be a random sample from:

$$f_y(y_i; \theta) = (\theta+1) y_i^{\theta} \qquad 0 < y < 1.$$

(4a) Find a method-of-moments estimator of θ.

(4b) Find the MLE of θ.

(4c) What is the asymptotic variance of the MLE of θ?

(4d) What is the sufficient statistic for θ?

(4e) Derive the distribution of the largest order statistic of a sample of size n from this distribution.

This completes the examination