MA 486/586-2D (Statistics), Dr Chernov Show your work.

Midterm test #1Thu, Feb 9, 2006

1. (20 pts) The following are test scores in a small class: 97 89 100 13 92 94 100 96 98 91

(a) Determine the mode, the median, the quartiles and the IQR.

(b) Draw a Box-and-Whisker diagram.

(c) Find the sample mean, the sample variance and the sample standard deviation.

(d) Which statistic – the sample mean, the median, or the mode – seems to better represent a typical score?

(Bonus) Remove the obvious outlier and recompute the same mean. Does it better represent a typical score?

2. (30 pts) A random sample of size n = 71 from $N(\mu, \sigma^2)$ yielded

$$\sum_{i=1}^{71} x_i = -142 \quad \text{and} \quad \sum_{i=1}^{71} x_i^2 = 564$$

(a) Compute a maximum likelihood estimate for μ . Did you use an unbiased estimate?

Answer: $\hat{\mu} = \bar{x} = -142/71 = 2$, it is unbiased.

- (b) Give an upper endpoint for a 90% confidence interval for μ .
- (c) Compute a maximum likelihood estimate for σ^2 .

Answer: V = 280/71.

(d) Compute an unbiased estimate for σ^2 .

Answer: $s^2 = 280/70 = 4$.

(e) Construct a 98% confidence interval for σ^2 using the χ^2 percentiles. (Bonus) Construct a 98% confidence interval for σ^2 using normal approximation.

3. (30 pts) Let x_1, \ldots, x_{11} and y_1, \ldots, y_{25} be two independent random samples from normal distributions $N(\mu_x, \sigma_x^2)$ and $N(\mu_y, \sigma_y^2)$, respectively. Their sample means are $\bar{x} = 6.4$ and $\bar{y} = -1.9$, and their sample variances are $s_x^2 = 33$ and $s_y^2 = 50$, respectively. (a) Construct a 90% confidence interval for the difference $\mu_x - \mu_y$. (Make sure that you are using the right assumptions!)

Hint: variances are unknown and unequal, use Welch's formula.

(b) Construct a 90% confidence interval for the ratio σ_x/σ_y .

4. (10 pts) Let p_M equal the proportion of adult males who regularly watch TV news and p_F equal the proportion of adult females who regularly watch TV news. A random poll of 1000 adult males showed that 425 of them regularly watch TV news, and a random poll of 800 adult females showed that 344 of them regularly watch TV news.

(a) Give point estimates for p_M and p_F .

(b) Give a point estimate for $p_M - p_F$.

(c) Construct a 96% confidence interval for $p_M - p_F$.

(Bonus) Give an upper endpoint for a 96% confidence interval for $p_M - p_F$. (Hint: use the main part of Table V to approximate the percentile $z_{0.04}$ the best you can.)

5. (10 pts) A botanist wants to estimate the proportion p of seeds (of a certain type) that germinate. She anticipates that at most 5% of them will fail to germinate. (a) How many seeds need be tested so that a 95% confidence interval for p shorter than 0.02 can be constructed?

Hint: $n \ge z_{\alpha/2}^2 \hat{p}(1-\hat{p})/\varepsilon^2$, where $\hat{p} = 0.95$.

(b) Suppose now the botanist has no anticipation of the proportion of germinating seeds. How many seeds does she need to test then, in order to construct a 95% confidence interval for p of length ≤ 0.02 ?

Hint: $n \ge z_{\alpha/2}^2/(4\varepsilon^2)$.