14.30 Exam 3 Solutions
Fall 2004
Bilal Zia

Question 1

(i) FALSE. Knowing the null distribution of a hypothesis test and the decision rule allows on to calculate α but not β.

(ii) FALSE. MLE estimators are consistent but not always unbiased. For instance, the MLE of the population variance is consistent but biased.

(iii) FALSE. We cannot make any generalized statements about MM estimators that are related to unbiasedness or consistency.

(iv) TRUE. This is just the invariance property of MLE.

(v) FALSE. Optimality implies minimizing β (Pr(Type II error)) for all possible values of the parameters in the alternative set, given the size of α. Hence, you cannot think of optimality as minimizing both α and β.

(vi) FALSE. The most efficient estimator is one with minimum variance. Minimum mean squared error is the minimum sum between variance and the estimator's bias squared.

(vii) FALSE. The sample mean has a normal distribution when the population distribution is known to be normal, or when the sample size is large enough that CLT will imply normality.

Question 2

a) $Y_i = \left\{ \begin{array}{ll} 0 & \text{if } X_i \leq h \\ 1 & \text{if } X_i > h \end{array} \right\}$; $h = 0.5$

pdf of discrete Y :

$\Pr(Y_i = 0) = \Pr(X_i \leq (h = 0.5))$

$= \int_0^{h=0.5} \frac{1}{\theta} dx = \frac{h}{\theta} = \frac{1}{2\theta}$

$\Pr(Y_i = 1) = \Pr(X_i > (h = 0.5))$

$= \int_{h=0.5}^{\theta} \frac{1}{\theta} dx = 1 - \frac{h}{\theta} = 1 - \frac{1}{2\theta}$

Therefore:

$f_Y(y) = \left\{ \begin{array}{ll} \frac{h}{\theta} & \text{if } y = 0 \\ 1 - \frac{h}{\theta} & \text{if } y = 1 \end{array} \right\}$; $h = 0.5$

which is a Bernoulli trial

(b) Method of Moments Estimator:

$E(Y) = 0 . \Pr(y = 0) + 1 . \Pr(y = 1)$

$= 1 - \frac{h}{\theta} = 1 - \frac{1}{2\theta}$

1st sample moment: $\frac{1}{n} \sum_{i=1}^{n} Y_i = \overline{Y}$

equating the moments:

$1 - \frac{h}{\theta} = \overline{Y}$

1

$\widehat{\theta}_{MM} = h.\frac{1}{1-\overline{Y}} = \frac{1}{2}.\frac{1}{1-\overline{Y}}$

(c) First step as always is to find the joint pdf of all the $Y_i s$

The Likelihood Function: $L(\theta \mid Y_1, Y_2, ..., Y_n) = \prod_{i=1}^{n} f_Y(y_i) = \left(\frac{h}{\theta}\right)^{n-\sum_{i=1}^{n} y_i} (1-$

$\frac{h}{\theta})^{\sum_{i=1}^{n} y_i}$

The Log-Likelihood Function: $(n-\sum_{i=1}^{n} y_i)(\ln h - \ln \theta) + \sum_{i=1}^{n} y_i(\ln(1-\frac{h}{\theta}))$

Now, maximizing the Log-Likelihood Function wrt θ and equating to 0 :

$\frac{n-\sum_{i=1}^{n} y_i}{\theta} = \frac{\sum_{i=1}^{n} y_i}{1-\frac{h}{\theta}}.\frac{h}{\theta^2}$

$\widehat{\theta}_{MLE} = \frac{hn}{n-\sum_{i=1}^{n} y_i} = h.\frac{1}{1-\overline{Y}} = \frac{1}{2}.\frac{1}{1-\overline{Y}}$

notice that $\widehat{\theta}_{MLE} = \widehat{\theta}_{MM}$

(d) The idea behind MLE is to maximize the likelihood function with respect to the unknown parameters. In part (c), we were given the value of h and therefore maximized the likelihood function with respect to unknown θ. Now, we are told that h is also unknown. The idea is the same with maximizing the likelihood function, but now we have to maximize over not only θ but also h.Mathematically:

Likelihood Function: $L(\theta, h \mid Y_1, Y_2, ..., Y_n)$

$\max_{\theta,h}[L(\theta, h \mid Y_1, Y_2, ..., Y_n)] = \max_{\theta,h}[\left(\frac{h}{\theta}\right)^{n-\sum_{i=1}^{n} y_i} (1 - \frac{h}{\theta})^{\sum_{i=1}^{n} y_i}]$

This maximization will give us our MLEs: $\widehat{\theta}_{MLE}$ and $\widehat{h}_{MLE}$

Question 3

(a) one-sided test:

$$H_0 \quad : \quad \mu = 600$$
$$H_1 \quad : \quad \mu > 600$$

Test statistic is the sample mean: $\overline{X} \sim N(600, \frac{60^2}{16})$ under the null hypothesis

Decision rule: Reject H_0 if $\overline{X} > k$, such that $\Pr(\overline{X} > k \mid \mu = 600) = 0.05$

$= \Pr(\overline{X} \leq k \mid \mu = 600) = 0.95$

$= \Pr(\frac{\sqrt{n}(\overline{X}-\mu)}{\sigma} \leq \frac{\sqrt{16}(k-600)}{60} \mid \mu = 600) = 0.95$

$= \Pr(\frac{\sqrt{n}(\overline{X}-\mu)}{\sigma} \leq \frac{\sqrt{16}(k-600)}{60} \mid \mu = 600) = 0.95$

$= \Pr(Z \leq 1.645) = 0.95$

$\frac{\sqrt{16}(k-600)}{60} = 1.645$

$k = 624.75$

Reject H_0 if:

$\overline{X} > 624.75$

(b) $u = 620$

Power$= 1 - \beta = 1 - \Pr(\text{don't reject } H_0 \mid H_1 \text{ is true})$

$= 1 - \Phi\left(\frac{624.75 - 620}{15}\right)$
$= 1 - \Phi(0.3167)$
$= 1 - 0.6255$
$= 0.3745$

Now if $u =: 640$
Power$= 1 - \Phi\left(\frac{624.75 - 640}{15}\right)$
$= 1 - \Phi(-1.016)$
$= \Phi(1.016)$
$= 0.8461$

Now if $u =: 680$
Power$= 1 - \Phi\left(\frac{624.75 - 680}{15}\right)$
$= 1 - \Phi(-3.683)$
$= \Phi(3.683)$
$= 0.9999$

Explanation: This question is meant to highlight the wording and intuition we use when testing hypotheses. When testing hypothesis, we should not think of "accepting" the null hypothesis if our results are such, but rather we should be thinking in terms of "failing to reject" the null hypothesis. Moreover, we should be thinking in terms of the probability of rejecting the null hypothesis, when in fact the alternative is true (i.e. $\Pr(\text{reject } H_0 \mid H_1) =$Power of test). We want this probability to be high for parameter values in the alternative set.

The exercise performed above is meant to highlight this distinction. As the probability of Type II error ($\Pr(\text{don't reject } H_0 \mid H_1) = \beta$) becomes larger, our probability of incorrectly accepting the null hypothesis becomes larger, and accordingly, the power of the test goes down. Therefore, we do not use the term "accept the null" as we might still be wrong, and instead, we use "cannot reject the null".

In terms of number of observations, our estimation becomes stronger as the size of the sample increases. Ideally, we would like to have the entire population as our sample, but that is in most cases unfeasible; that is why we take samples of the population to test predictions concerning the population. The larger the sample we can get, the more accurate our results will portray the true population, and the greater will be the power of our test.

(c) Since variance is unknown, we first have to estimate the population variance σ^2 with the sample variance s^2. Given this estimation, we have to use the t-distribution with the following test statistic:

Test statistic: $\frac{\overline{X} - \mu}{\frac{s}{\sqrt{n}}} \backsim t_{16-1}$

reject H_0 if $\frac{\overline{X} - \mu}{\frac{s}{\sqrt{n}}} < t_{15}(\alpha)$

where α is the significance level.

(d) Now we are looking for a 2-sided test which has the same alternative hypothesis as in part (a). So obviously we need to adjust our null hypothesis:

$$H_0 \quad : \qquad \mu \leq 600$$
$$H_1 \quad : \qquad \mu > 600$$

Decision rule: Reject H_0 if $\overline{X} > k$, such that $\Pr(\overline{X} > k \mid \mu \leq 600) = \alpha$

This means that

$\Pr(\overline{X} \leq k \mid \mu) = 1 - \alpha$

$= \Pr(\frac{\sqrt{n}(\overline{X}-\mu)}{\sigma} \leq \frac{\sqrt{16}(k-\mu)}{60} \mid \mu) = 1 - \alpha$

$= \Pr(\frac{\sqrt{n}(\overline{X}-\mu)}{\sigma} \leq \frac{\sqrt{16}(k-\mu)}{60} \mid \mu) = 1 - \alpha$

$= \Pr(Z \leq \frac{\sqrt{16}(k-\mu)}{60}) = 1 - \alpha$

$\Phi\left(\frac{\sqrt{16}(k-\mu)}{60}\right) = (1 - \alpha)$

now as $\mu \downarrow \rightarrow \Phi\left(\frac{\sqrt{16}(k-\mu)}{60}\right) \uparrow \rightarrow (1 - \alpha) \uparrow \rightarrow \alpha \downarrow$

So, the $\sup_{\mu}\alpha_\mu = \alpha_{\mu=600} = 0.05 = 5\%$

Question 4

(a) Two-sided hypothesis test:

$$H_0 \quad : \qquad \mu = 0.5$$
$$H_1 \quad : \qquad \mu \neq 0.5$$

Decision Rule: Reject H_0 if $\overline{X} \notin [C_1, C_2]$

such that:

$\Pr(\overline{X} < C_1 \mid \mu = 0.5) = 0.015$; $\Pr(\overline{X} > C_2 \mid \mu = 0.5) = 0.015$

Since n is large, we can apply CLT and use either the z tables or t tables as they will give identical values when n is large.

Using the same techniques as in Question 3 part (a), we get:

$C_1 = 0.5 - 2.17 \cdot \frac{0.3}{20} = 0.46745$

$C_2 = 0.5 + 2.17 \cdot \frac{0.3}{20} = 0.53255$

Summarizing, reject null hypothesis if $\overline{X} > 0.53255$ or $\overline{X} < 0.46745$

(b) Since $\overline{X} = 0.539 > 0.53255$, we reject the null hypothesis

P-value$= \Pr(|\overline{X}| > 0.539 \mid \mu = 0.5)$

$= \Pr(\left|\frac{\sqrt{n}(\overline{X}-\mu)}{\sigma}\right| > \frac{20(0.539-0.5)}{0.3} \mid \mu = 0.5)$

$= \Pr(\left|\frac{\sqrt{n}(\overline{X}-\mu)}{\sigma}\right| > 2.6 \mid \mu = 0.5)$

$= \Pr(|Z| > 2.6 \mid \mu = 0.5)$

$= 2\Pr(Z > 2.6 \mid \mu = 0.5)$

$= 2(1 - \Pr(Z < 2.6 \mid \mu = 0.5))$

$$= 2(1 - 0.9953)$$
$$= 0.0094$$

(c) In part (a) the sample size was large so we invoked the CLT and assumed a normal distribution. Hence, by using CLT, we are saying the distribution is approximately normal. It is important to note, however, that CLT applies as n goes to infinity. For n smaller than infinity, it is an approximation; of course the larger the n, the better the approximation. The fact that we know σ, the population variance, is not significant in deciding whether to use normal or t distributions. Moreover, even if population variance was unknown, we could still have used a normal distribution via the distributional implications of CLT (law of large numbers implies $s \xrightarrow{p} \sigma$).

Rather, the choice between t and normal distributions is affected by our desire to be conservative. In general, the t-distribution is "fatter" than the normal distribution. For smaller sample sizes, using the t-distribtuion is a conservative approach, as we are less likely to reject the null if we use the t-distribution rather than the normal due to the added uncertaintly introduced through estimating σ^2. Hence, the changes that one would make from part (a) would be to use the t-dist instead of normal, and perhaps also reduce the size of the test (use a lower α).