Economics 300–01: Quantitative Methods in Economics
Wesleyan University, Spring 2002

Summary of Formulas for Midterm #2

1. **Sampling Distributions:**

   (a) Single Random Sample

   - If $X \sim N(\mu,\, \sigma^2)$, then $\bar{X} \sim N(\mu,\, \frac{\sigma^2}{n})$.
   - If $X \sim \ (\mu,\, \sigma^2)$, then $\bar{X} \dot{\sim} N(\mu,\, \frac{\sigma^2}{n})$.
   - If $X \sim B(n,\, \pi)$, then $P \dot{\sim} N(\pi,\, \frac{\pi(1-\pi)}{n})$.

   (b) Two Independent Random Samples

   - If $X \sim N(\mu_X,\, \sigma_X^2)$ and $Y \sim N(\mu_Y,\, \sigma_Y^2)$, then $\bar{X} - \bar{Y} \sim N\left(\mu_X - \mu_Y,\, \frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}\right)$.
   - If $X \sim \ (\mu_X,\, \sigma_X^2)$ and $Y \sim \ (\mu_Y,\, \sigma_Y^2)$, then $\bar{X} - \bar{Y} \dot{\sim} N\left(\mu_X - \mu_Y,\, \frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}\right)$.
   - If $X \sim B(n_X,\, \pi_X)$ and $Y \sim B(n_Y,\, \pi_Y)$,
     then $P_X - P_Y \dot{\sim} N\left(\pi_X - \pi_Y,\, \frac{\pi_X(1-\pi_X)}{n_X} + \frac{\pi_Y(1-\pi_Y)}{n_Y}\right)$.

2. **Properties of Estimators:**

   (a) Unbiased: $\hat{\theta}$ is unbiased if $\mathrm{E}[\hat{\theta}] = \theta$.

   - $\mathrm{bias}(\theta) = \mathrm{E}[\hat{\theta}] - \theta$

   (b) Efficiency: $\hat{\theta}_1$ is more efficient than $\hat{\theta}_2$ if $\mathrm{MSE}(\hat{\theta}_1) < \mathrm{MSE}(\hat{\theta}_2)$.

   - $\mathrm{MSE}(\hat{\theta}) = \mathrm{E}[(\hat{\theta} - \theta)^2] = \mathrm{bias}(\hat{\theta})^2 + \mathrm{var}(\hat{\theta})$

   (c) Consistency: $\hat{\theta}$ is consistent if, as $n \to \infty$, $\mathrm{MSE}(\hat{\theta}) \to 0$.
   (*Note:* This is a sufficient condition, not a necessary one.)

   (d) Best Unbiased Estimator (BUE):
   $\hat{\theta}$ is the BUE if, for any other unbiased estimator $\tilde{\theta}$, $\mathrm{var}(\hat{\theta}) \leq \mathrm{var}(\tilde{\theta})$.

3. **Confidence Intervals: Single Sample**

   (a) General form: The $100(1-\alpha)\%$ confidence interval for θ is

   - $\hat{\theta} \pm c_{\alpha/2}\, SE_{\hat{\theta}}$

   (b) $X \sim N(\mu,\, \sigma^2)$: The $100(1-\alpha)\%$ confidence interval for μ is

   - $\bar{X} \pm z_{\alpha/2}\, \frac{\sigma}{\sqrt{n}}$, if σ^2 known (exact)
   - $\bar{X} \pm t_{\alpha/2,\, n-1}\, \frac{s}{\sqrt{n}}$, if σ^2 unknown (exact)

(c) $X \sim\ (\mu,\, \sigma^2)$: The $100(1-\alpha)\%$ confidence interval for μ is

- $\bar{X} \pm z_{\alpha/2}\, \frac{\sigma}{\sqrt{n}}$, if σ^2 known (approximate)
- $\bar{X} \pm z_{\alpha/2}\, \frac{s}{\sqrt{n}}$, if σ^2 unknown (approximate)

(d) $X \sim B(n,\, \pi)$: The $100(1-\alpha)\%$ confidence interval for π is

- $P \pm z_{\alpha/2} \sqrt{\frac{P(1-P)}{n}}$ (approximate)

4. **Confidence Intervals: Two Independent Samples**

(a) If $X \sim (\mu_X,\, \sigma_X^2)$ and $Y \sim (\mu_Y,\, \sigma_Y^2)$,
the $100(1-\alpha)\%$ confidence interval for $\mu_X - \mu_Y$ is:

- Case 1: $\sigma_X^2 = \sigma_Y^2 = \sigma^2$, known

$$\bar{X} - \bar{Y} \pm z_{\alpha/2}\, \sigma \sqrt{\frac{1}{n_X} + \frac{1}{n_Y}}$$

- Case 2: $\sigma_X^2 = \sigma_Y^2 = \sigma^2$, unknown

$$\bar{X} - \bar{Y} \pm t_{\alpha/2,\,(n_X+n_Y-2)}\, s_p \sqrt{\frac{1}{n_X} + \frac{1}{n_Y}}, \quad \text{where } s_p = \sqrt{\frac{(n_X-1)\, s_X^2 + (n_Y-1)\, s_Y^2}{n_X + n_Y - 2}}$$

- Case 3: $\sigma_X^2 \neq \sigma_Y^2$, known

$$\bar{X} - \bar{Y} \pm z_{\alpha/2} \sqrt{\frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}}$$

- Case 4: $\sigma_X^2 \neq \sigma_Y^2$, unknown

$$\bar{X} - \bar{Y} \pm z_{\alpha/2} \sqrt{\frac{s_X^2}{n_X} + \frac{s_Y^2}{n_Y}}$$

(*Note:* The text uses the t distribution in this case when n is small.)

(b) If $X \sim B(\mu_X,\, \sigma_X^2)$ and $Y \sim B(\mu_Y,\, \sigma_Y^2)$,
the $100(1-\alpha)\%$ confidence interval for $\pi_X - \pi_Y$ (approximate) is:

$$P_X - P_Y \pm z_{\alpha/2} \sqrt{\frac{P_X(1-P_X)}{n_X} + \frac{P_Y(1-P_Y)}{n_Y}}$$

5. **Confidence Intervals: Matched (Paired) Samples**

- Let X_1 be the data from sample 1, X_2 be the data from sample 2.
  Let n be the sample size. Define:
  - $\Delta = \mu_1 - \mu_2$
  - $D_i = X_{1i} - X_{2i}$ for all i
  - $\overline{D} = \bar{X}_1 - \bar{X}_2$
  - $s_D^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(D_i - \overline{D}\right)^2$
- Pivotal Quantity:
  $$\frac{\overline{D} - \Delta}{s_D/\sqrt{n}} \sim t_{\alpha/2,\, n-1}$$
- The $100(1-\alpha)\%$ confidence interval for Δ is $\overline{D} \pm t_{\alpha/2,\, n-1} \frac{s_D}{\sqrt{n}}$

6. **Hypothesis Testing:**

   (a) Two-sided alternative

   - Null Hypothesis: $\mathrm{H}_0 : \theta = \theta_0$. Alternative Hypothesis: $\mathrm{H}_1 : \theta \neq \theta_0$.
   - Classical Decision Rule:
     $$\text{Reject } \mathrm{H}_0 \text{ at the } \alpha \text{ level of significance if } \left| \frac{\hat{\theta} - \theta_0}{SE_{\hat{\theta}}} \right| > c_{\alpha/2}\,.$$
   - Decision Rule using Confidence Interval:
     $$\text{Reject } \mathrm{H}_0 \text{ at the } \alpha \text{ level of significance if } \theta_0 \text{ lies outside } \hat{\theta} \pm c_{\alpha/2}\, SE_{\hat{\theta}}\,.$$
   - p-value:
     $$p = 2 \cdot \Pr\left(C > \left| \frac{\hat{\theta} - \theta_0}{SE_{\hat{\theta}}} \right| \right)$$
     where C is a random variable with the same distribution as the test statistic. (For the problems above, either a t or z distribution.)

   (b) One-sided alternative (greater)

   - Null Hypothesis: $\mathrm{H}_0 : \theta = \theta_0$. Alternative Hypothesis: $\mathrm{H}_1 : \theta > \theta_0$.

- Classical Decision Rule:

$$\text{Reject H}_0 \text{ at the } \alpha \text{ level of significance if } \left(\frac{\hat{\theta} - \theta_0}{SE_{\hat{\theta}}}\right) > c_\alpha \,.$$

- Decision Rule using Confidence Interval:

$$\text{Reject H}_0 \text{ at the } \alpha \text{ level of significance if } \theta_0 \text{ is less than } \hat{\theta} - c_\alpha\, SE_{\hat{\theta}} \,.$$
$$\text{(Equivalently, reject H}_0 \text{ if } \hat{\theta} \text{ is greater than } \theta_0 + c_\alpha\, SE_{\hat{\theta}} \,.)$$

- p-value:

$$p = \Pr\left(C > \frac{\hat{\theta} - \theta_0}{SE_{\hat{\theta}}}\right)$$

where C is a random variable with the same distribution as the test statistic. (For the problems above, either a t or z distribution.)

(c) One-sided alternative (lesser)

- Null Hypothesis: $\text{H}_0 : \ \theta = \theta_0$. Alternative Hypothesis: $\text{H}_1 : \ \theta < \theta_0$.
- Classical Decision Rule:

$$\text{Reject H}_0 \text{ at the } \alpha \text{ level of significance if } \left(\frac{\hat{\theta} - \theta_0}{SE_{\hat{\theta}}}\right) < c_\alpha \,.$$

- Decision Rule using Confidence Interval:

$$\text{Reject H}_0 \text{ at the } \alpha \text{ level of significance if } \theta_0 \text{ is greater than } \hat{\theta} + c_\alpha\, SE_{\hat{\theta}} \,.$$
$$\text{(Equivalently, reject H}_0 \text{ if } \hat{\theta} \text{ is less than } \theta_0 - c_\alpha\, SE_{\hat{\theta}} \,.)$$

- p-value:

$$p = \Pr\left(C < \frac{\hat{\theta} - \theta_0}{SE_{\hat{\theta}}}\right)$$

where C is a random variable with the same distribution as the test statistic. (For the problems above, either a t or z distribution.)

7. **Type I and Type II Errors:**

(a) $\Pr(\text{Type I error}) = \Pr\big(\text{Reject H}_0 \,\big|\, \text{H}_0 \text{ true}\big) = \alpha$

- $\alpha =$ "size" of a test

(b) $\Pr(\text{Type II error}) = \Pr\big(\text{Do not reject H}_0 \,\big|\, \text{H}_1 \text{ true}\big) = \beta$

- $1 - \beta =$ "power" of a test

Revised: 26-Mar-2002