17.800 Midterm Review

Yiqing Xu

1 Probability

- Probability rules
- Conditional probability
- Bayes' rule

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)}$$

- Independence
- Conditional independence

2 Random Variables

- PDF, PMF, CDF
- Expectation

$$E[X] = \sum_x xP_X(x)$$

- Variance

$$\sigma_X^2 = E[(X - E[X])^2] = E[X^2] - E[X]^2$$

- Normal distribution
- Conditional expectation
- Conditional variance
- Covariance

$$Cov[X, Y] = E[XY] - E[X]E[Y]$$

- Correlation

$$Cor[X, Y] = \frac{Cov[X, Y]}{\sqrt{V[X]V[Y]}}$$

3 Point Estimation

- Estimands, estimators, and estimates
- Unbiasedness
- Efficiency
- Consistency

- Asymptotic normality (def.)
- Mean squared error

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

- Law of large numbers (def.)
- Central limit theorem (def.)

4 Interval Estimation

- Estimating population variance

$$\hat{\sigma}^2 \equiv S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$$

- Standard error of $\bar{X}_n$

$$SE(\bar{X}_n) = S_n/\sqrt{n}$$

- Confidence interval

$$[\bar{X}_n - 1.96SE(\bar{X}), \bar{X}_n + 1.96SE(\bar{X}_n)]$$

- t-distribution and t-statistic

$$T_n = \frac{\bar{X}_n - \mu}{S_n/\sqrt{n}} = \frac{\bar{X}_n - \mu}{SE(\bar{X}_n)} \sim \tau_{n-1}$$

5 Hypothesis Testing

- Type I error (significance level)
- Type II error (red power)
- One-sample t-test

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim \tau_{n-1}$$

- Two-sample t-test

$$T = \frac{\bar{Y}_A - \bar{Y}_B}{\sqrt{S_A^2/N_A + S_B^2/N_B}} \sim^{approx.} N(0,1)$$

- p-values
- Hypothesis testing