Agendas

Lecture – Quiz #1 (9/17/18)

Lecture – Quiz #1 Review (9/19/18)

Lecture – Ch 3 Lecture 1 - Slides 1 – 14 (9/21/18)

1. Random Variables – numerical representation of an outcome of a random experiment
2. Discrete vs Continuous Random Variables (the support of the RV & how to calculate probabilities)
3. Discrete Random Variables – pdf, 2 key properties of all pdfs
4. Forms of Discrete Random Variables – Table (Info Given), Probability Histogram, Formula

Lecture – Ch 3 Lecture 2 Slides 15 – 25 (9/24/18)

1. Expected Value – Definition, Constant Rule, Function Rule (with example), $E[X^2 + 3X + 5] = E[X^2] + 3E[X] + 5$, and $E[c] = c$
2. Variance & Standard Deviation – Definition as $E[(X - \mu)^2]$, shortcut formula (with proof), $V[X] = a^2V[X]$
3. Bernoulli Distribution $p(x) = p^x(1 - p)^{1-x}$ $x = 0,1$, 1 trial with 2 possible outcomes (coin flip, but don't have to have equal probability of success and failure), $X \sim Bern(p)$, proofs of expectation and variance $E[X] = p$, $V[X] = p(1 - p)$
4. Binomial Distribution – extension of Bernoulli distribution, FITS, proof of mean $E[X] = np$, just state the variance $V[X] = np(1 - p)$

Lecture – Ch 3 Lecture 3 Slides 26 – 34 (9/26/18)

1. Show that if $Y \sim Bin(n, p)$ and $Z \sim Bin(n, 1 - p)$, then $p(y = n - y) = P(z = y)$.
2. Consider the RV $Y \sim Bin(n, p)$. Recursively define the $p(y)$.
3. Derive the variance of any RV $Y \sim Bin(n, p)$
4. Geometric Distribution
   a. Context
   b. Proof of Valid pdf
   c. Derivation of expectation

   d. The memoryless property for the Geometric distribution states that if $X \sim Geo(p)$, with i, j positive numbers, then $P(X \geq i + j | X \geq i) = P(X \geq j)$.
      i. First, determine the cumulative distribution function for a geometric distribution (hint: the cumulative distribution function $F(x) = P(X \leq x) = P(you\ get\ a\ success\ at\ or\ before\ the\ x^{th} trial) = 1 - P(first\ x\ trials\ are\ all\ failures)$)
      ii. Now, prove that the Geometric distribution indeed exhibits the memoryless property.
5. Hypergeometric Distribution
   a. Context
   b. Proof of valid pdf

$$X \sim Geo(p)$$
$$p(x) = (1-p)^{x-1} p \quad x = 1,2,3, \ldots$$
$$E[X] = \frac{1}{p} \quad V[X] = \frac{1-p}{p^2}$$

$$X \sim Hypergeo(N, M, n)$$
$$p(x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}} \quad x = 0,1,2,3, \ldots n \quad n \leq M \leq N \quad n \leq N - M$$
$$E[X] = \frac{nM}{N} \quad V[X] = n\left(\frac{M}{N}\right)\left(\frac{N-M}{N}\right)\left(\frac{N-n}{N-1}\right)$$

Lecture – Ch 3 Lecture 4 Slides 35 – 45 (9/28/18)

1. Geometric Distribution
   a. Context
   b. Proof of Valid pdf
   c. Derivation of expectation
   d. The memoryless property for the Geometric distribution states that if $X \sim Geo(p)$, with i, j positive numbers, then $P(X \geq i + j | X \geq i) = P(X \geq j)$.
      i. First, determine the cumulative distribution function for a geometric distribution (hint: the cumulative distribution function $F(x) = P(X \leq x) = P(you\ get\ a\ success\ at\ or\ before\ the\ x^{th} trial) = 1 - P(first\ x\ trials\ are\ all\ failures)$)
      ii. Now, prove that the Geometric distribution indeed exhibits the memoryless property.
2. Hypergeometric Distribution
   a. Context
   b. Proof of valid pdf
3. Negative Binomial Distribution

    a. Context
4. Poisson Distribution
    a. Proof of valid pdf
    b. Proof of mean & variance

$$X \sim Geo(p)$$
$$p(x) = (1-p)^{x-1}p \quad x = 1,2,3,\ldots$$
$$E[X] = \frac{1}{p} \quad V[X] = \frac{1-p}{p^2}$$

$$X \sim Hypergeo(N, M, n)$$
$$p(x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}} \quad x = 0,1,2,3,\ldots n \quad n \le M \le N \quad n \le N-M$$
$$E[X] = \frac{nM}{N} \quad V[X] = n\left(\frac{M}{N}\right)\left(\frac{N-M}{N}\right)\left(\frac{N-n}{N-1}\right)$$

$$X \sim Negbin(r, p)$$
$$p(x) = \binom{x-1}{r-1}p^r(1-p)^{x-r} \quad x = r, r+1, r+2, \ldots \quad 0 \le p \le 1$$
$$E[X] = \frac{r}{p} \quad V[X] = \frac{r(1-p)}{p^2}$$

$$X \sim Poisson(\lambda)$$
$$p(x) = \frac{e^{-\lambda}\lambda^x}{x!} \quad x = 0,1,2,\ldots \quad 0 < \lambda$$

$$E[X] = V[X] = \lambda$$

Lecture – Problem Session for Ch 3 (10/1/18)

Lecture – Ch 4 Lecture 1 Slides 1 – 5 (10/3/18)

1. Finish Review Question 4
2. Probability Density Functions, Probability Distribution Functions, and Probability Mass Functions
3. 2 Distinguishing properties of continuous random variables – support is an interval & $P(X = x) = 0$ for any singular value of x in the support.
4. Finding probabilities using a pdf ($P(a \le X \le b) = \int_a^b f(x)dx$)
5. 2 Properties of every legitimate pdf $f(x) \ge 0 \ \forall x \ \epsilon \ \chi$ & $\int_{-\infty}^{\infty} f(x) = 1$
6. Verify whether the following function is a valid probability density function on the given support

$$f(x) = \begin{cases} .075x + .2 & 3 \le x \le 5 \\ 0 & otherwise \end{cases}$$

7. Verify whether the following function is a valid pdf on the given support

$$f(x) = \begin{cases} x - 4 & 1 \le x \le 4 - \sqrt{11} \\ 0 & otherwise \end{cases}$$

8. Verify whether the following function is a valid pdf on the given support

$$f(x) = \begin{cases} \frac{1}{25}y & 0 \le y < 5 \\ \frac{2}{5} - \frac{1}{25}y & 5 \le y \le 10 \\ 0 & y < 0 \text{ } or \text{ } y > 10 \end{cases}$$

9. Consider the function $f(x) = \begin{cases} kx^3 + \frac{k}{2}x^2 & 0 \le x \le 2 \\ 0 & otherwise \end{cases}$
    a. Find k such that the above function is a valid probability density function on the given support.
    b. What is the probability that $x \in (0,1)$? Is this the same as the probability $x \in [0,1]$? Why is this so?

Lecture – Ch 4 Lecture 2 Slides 6 – 19 (10/5/18)

1. Uniform Distribution – Context, pdf ($\int_A^B c\, dx$)
2. CDF $F(x) = P(X \le x) = \int_{-\infty}^{x} f(y)dy$ (for $X \sim U(a,b)$ $F(x) = \frac{x-a}{b-a}$)
3. Percentiles
4. Expectation and Variance -> Go through uniform distribution ($E[X] = \frac{a+b}{2}$ $V[X] = \frac{(b-a)^2}{12}$)
5. Suppose that Y has a uniform distribution over the interval [0,1].
    a. Prove that the distribution of Y is valid
    b. Find $F(y)$
    c. Show that $P(a \le Y \le a + b)$, for $a \ge 0, b \ge 0$, and $a + b \le 1$ depends only upon the value of b.
    d. What is the third moment, i.e. $E(X^3)$?
    e. What is $E[X^3 - X^2 - 1]$?
6. A telephone call arrived at a switchboard at random within a 1-minute interval. The switch board was fully busy for 15 seconds into this 1-minute period. What is the probability that the call arrived when the switchboard was not fully busy?
7. Consider the following pdf

$$f(x) = \begin{cases} x & 0 \le x \le 1 \\ \frac{1}{2} & 1 < x \le 2 \\ 0 & o.w. \end{cases}$$

    a. Is the distribution differentiable everywhere? How about continuous?

b. Find $F(x)$

8. The cycle time for trucks hauling concrete to a highway construction site is uniformly distributed over the interval 50 to 70 minutes.
    a. What is the probability that the cycle time exceeds 65 minutes if it is known that the cycle time exceeds 55 minutes?
    b. Calculate the IQR of this dataset

Lecture – Ch 4 Lecture 3 Slides 20 – 28 (10/8/18)

1. The cycle time for trucks hauling concrete to a highway construction site is uniformly distributed over the interval 50 to 70 minutes.
    a. What is the probability that the cycle time exceeds 65 minutes if it is known that the cycle time exceeds 55 minutes?
    b. Calculate the IQR of this dataset
2. Normal Distribution $\left(f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}\right)$
3. Standardization
4. Empirical Rule
5. Let Y be normally distributed with mean 4 and variance 1. Find the following:
    a. The range of values seen as "typical"
    b. The 84th percentile
    c. What percentage of observations fall below 2?
    d. What percentage of points fall above 5?
6. Let X be normally distributed with mean 10 and standard deviation 5. Find the following probabilities:
    a. P(X ≤ 8)
    b. P(X ≥ 11)
    c. The 86th percentile
    d. Q1 of this distribution
    e. The median of this distribution
    f. The value such that 16% of the data fall above this point.
    g. P(1 ≤ |X|)
7. Binomial Approximation – Conditions ($np \geq 10$ and $n(1-p) \geq 10$) & Distribution $\left(X \sim N\big(np, np(1-p)\big)\right)$
8. Continuity Correction – using a continuous RV to approximate a discrete random variable $\big(P(X \geq a) = P(X \geq a - 0.5)\big)$
9. Suppose that 10% of all steel shafts produced by a certain process are nonconforming but can be reworked (rather than having to be scrapped). Consider a random sample of 200 shafts, and let X denote the number among these that are nonconforming and can be reworked. What is the (approximate) probability that X is
    a. At most 30?
    b. Less than 30?
    c. Between 15 and 25 (inclusive)?

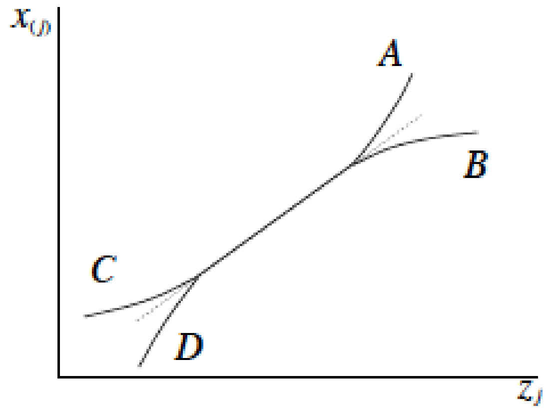
Lecture – Ch 4 Lecture 4 Slides 29 – 36 (10/10/18)

1. Binomial Approximation – Conditions ($np \geq 10$ and $n(1-p) \geq 10$) & Distribution $\left(X \sim N\left(np, np(1-p)\right)\right)$
2. Continuity Correction – using a continuous RV to approximate a discrete random variable $\left(P(X \geq a) = P(X \geq a - 0.5)\right)$
   a. The idea: Maximize the area that you are considering under the normal distribution such that the inequality still represents the same possible values of X under the binomial distribution.
      i. $P(X \leq 120) = P(X \leq 120.5)$
      ii. $P(X < 120) = P(X \leq 119) = P(X \leq 119.5)$
      iii. $P(130 \leq X \leq 145) = P(129.5 \leq X \leq 145.5)$
3. Suppose that 10% of all steel shafts produced by a certain process are nonconforming but can be reworked (rather than having to be scrapped). Consider a random sample of 200 shafts, and let X denote the number among these that are nonconforming and can be reworked. What is the (approximate) probability that X is
   a. At most 30?
   b. Less than 30?
   c. Between 15 and 25 (inclusive)?
4. Gamma Function & Its Properties
   a. $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$ (for $\alpha > 0$)
   b. For any $\alpha > 1, \Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$
   c. For any positive integer *n*, $\Gamma(n) = (n-1)!$
   d. $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$
   e. Evaluate the following:
      i. $\int_0^\infty x^3 e^{-x} dx$
      ii. $\int_0^\infty x e^{-\frac{1}{3}x} dx$
      iii. $\int_0^\infty x^2 e^{-\frac{1}{3}x} dx$
      iv. $\int_0^\infty x e^{-\frac{1}{3}x^2} dx$
5. Gamma Distribution $\left(X \sim Gamma(\alpha, \beta)\right)$
   a. Proof of valid pdf $\left(f(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}} \quad x \geq 0\right)$ with $\alpha, \beta > 0$
   b. Derivation of expectation and Variance ($E(X) = \alpha\beta$ and $V(X) = \alpha\beta^2$)
6. Exponential Distribution $\left(X \sim Exp(\lambda)\right)$
   a. pdf $\left(f(x) = \lambda e^{-\lambda x} \quad x \geq 0\right)$ with $\lambda > 0$ (If $X \sim Gamma(1, \beta)$, look familiar?)
   b. cdf $\left(F(x) = 1 - e^{-\lambda x} \quad x \geq 0\right)$
   c. Expectation and Variance ($E(X) = \frac{1}{\lambda}$ and $V(X) = \frac{1}{\lambda^2}$)
   d. Memoryless Property $P(X \geq t + t_0 | X \geq t_0) = P(X \geq t) = e^{-\lambda t}$
7. Chi-Squared Distribution $\left(X \sim \chi^2(\nu)\right)$

a. pdf $\left(f(x) = \frac{1}{2^{\nu/2}\Gamma\left(\frac{\nu}{2}\right)} x^{\left(\frac{\nu}{2}\right)-1} e^{-\frac{x}{2}} \quad x \geq 0\right)$ with $\nu = \text{d.f.}$ (If $X \sim Gamma(\frac{\nu}{2}, 2)$, look familiar?)
b. Expectation and Variance $(E(X) = \nu$ and $V(X) = 2\nu)$

Lecture – Ch 4 Lecture 5 Slides 37 – 46 (10/12/18)

1. Gamma Distribution $\left(X \sim Gamma(\alpha, \beta)\right)$
   a. Proof of valid pdf $\left(f(x) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}} \quad x \geq 0\right)$ with $\alpha, \beta > 0$
   b. Derivation of expectation and Variance $(E(X) = \alpha\beta$ and $V(X) = \alpha\beta^2)$
2. Chi-Squared Distribution $\left(X \sim \chi^2(\nu)\right)$
   a. pdf $\left(f(x) = \frac{1}{2^{\nu/2}\Gamma\left(\frac{\nu}{2}\right)} x^{\left(\frac{\nu}{2}\right)-1} e^{-\frac{x}{2}} \quad x \geq 0\right)$ with $\nu = \text{d.f.}$ (If $X \sim Gamma(\frac{\nu}{2}, 2)$, look familiar?)
   b. Expectation and Variance $(E(X) = \nu$ and $V(X) = 2\nu)$
3. Exponential Distribution $\left(X \sim Exp(\lambda)\right)$
   a. pdf $\left(f(x) = \lambda e^{-\lambda x} \quad x \geq 0\right)$ with $\lambda > 0$ (If $X \sim Gamma\left(1, \frac{1}{\lambda}\right)$, look familiar?)
   b. cdf $\left(F(x) = 1 - e^{-\lambda x} \quad x \geq 0\right)$
   c. Expectation and Variance $(E(X) = \frac{1}{\lambda}$ and $V(X) = \frac{1}{\lambda^2})$
   d. Memoryless Property $P(X \geq t + t_0 | X \geq t_0) = P(X \geq t) = e^{-\lambda t}$
4. Probability Plots
   a. Motivation -> Check if sample came from a certain type of distribution
   b. Creating the Q-Q (quantile-quantile plot)
      i. Sort the data from smallest to largest
      ii. For $i = 1,2, \dots, n$, $x_{(i)}$ is the $p_i = \frac{100(i-0.5)}{n} th$ sample percentile
      iii. Then calculate the theoretical percentiles of a standard normal distribution for each percentile $(P(z \leq z_i) = p_i)$
      iv. Plot the z_i on the x-axis and the $x_{(i)}$ on the y-axis
   c. Exercise: Suppose a data set contains 10 observations that are sorted as follows:
      -1.91, -1.25, -0.75, -0.53, 0.20, 0.35, 0.72, 0.87, 1.40, 1.56
      i. Construct the Normality Plot (aka the Q-Q Plot)
      ii. Describe the shape of the distribution
   d. Interpretation



i. Ideally on the 45° line → implies normality
ii. If the right tail bends upward, then distribution of the observations has a heavy right tail
iii. If the right tail bends downward, then distribution of the observations has a light right tail
iv. If the left tail bends upward, then distribution of the observations has a light left tail
v. If the left tail bends downward, then distribution of the observations has a heavy left tail

5. Beta Distribution
    a. Beta Function $\left(\int_0^1 x^{a-1}(1-x)^{b-1}dx = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}\right)$
    b. Proof of valid pdf $\left(f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} \;\; 0 \leq \; x \leq 1\right)$ with $\alpha, \beta > 0$
    c. Expectation and Variance $(E(X) = \frac{\alpha}{\alpha+\beta}$ and $V(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)})$
6. (Weibull Distribution)
7. (Lognormal Distribution)

Lecture – Problem Session for Ch 4 (10/15/18)

Lecture – Quiz #2 (10/17/18)

Lecture – Quiz #2 Review (10/19/18)

Lecture – Ch 6 Lecture 1 Slides 1 – 11 (10/22/18)

1. Beta Distribution
    a. Beta Function $\left(\int_0^1 x^{a-1}(1-x)^{b-1}dx = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}\right)$
    b. Proof of valid pdf $\left(f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} \;\; 0 \leq \; x \leq 1\right)$ with $\alpha, \beta > 0$

   c. Expectation and Variance ($E(X) = \frac{\alpha}{\alpha+\beta}$ and $V(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$)
2. General Introduction to Inference (Statistics Lifecycle)
3. Point Estimator versus Estimate (Definition of a Statistic)
4. Unbiased Estimator $\left(E(\hat{\theta}) = \theta\right)$ and bias (the difference)
5. Example 1 – When X is a binomial RV with parameters n and p, is the sample proportion $\hat{p} = \frac{X}{n}$ an unbiased estimator of p?
6. Mean Squared Error (MSE of an estimator = $E\left[\left(\hat{\theta} - \theta\right)^2\right]$
7. What happens to MSE when we have an unbiased estimator?
8. Example 1 Continued – Find the MSE of the estimator X/n and show that as $p \to 0$ the MSE of the estimator approaches 0.
9. Example 2 - The reading on a voltage meter connected to a test circuit is uniformly distributed over the interval $(\theta, \theta + 1)$, where θ is the true but unknown voltage of the circuit. Let $X_1, X_2, \dots, X_n$ be a random sample of readings from this voltage meter.
   a. Show that $\bar{X}$ is a biased estimator of θ and find its bias
   b. Is the estimator $X_{(n)} - X_{(1)}$ unbiased?
   c. Which is a better estimator, $X_{(n)} - X_{(1)}$ or $\bar{X}$?
   d. Can we find a better estimator of θ?

Lecture – Ch 6 Lecture 2 Slides 12 – 28 (10/24/18)

1. Measuring precision - standard error of an estimator $\left(\sigma_{\hat{\theta}} = \sqrt{V(\hat{\theta})}\right)$
2. Minimum Variance Unbiased Estimator (MVUE)
3. Example 1 (Cont.) – Quantify the precision of the estimator $\hat{p}_1 = \frac{X}{n}$. Which is a better estimator, $\hat{p}_1 = \frac{X}{n}$ or $\hat{p}_2 = \frac{X_1+X_2}{2n}$ (assume $X_1, \dots, X_n$ are i.i.d)?
4. Method of Moments
   a. What are moments ($\eta_r = E[X^r]$) Gamma Distribution
   b. How to find estimators using the method of moments (Set

$$\eta_1 = \frac{1}{n}\sum_{i=1}^{n} X_i$$
$$\eta_2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$$
$$\dots$$
$$\eta_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$$

   Then, isolate for the parameters you are estimating
   c. Prove that the moment estimators for $N(\mu, \sigma^2)$ are $\hat{\mu} = \bar{X}$ and $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$

d. Prove the moment estimators for $Gamma(\alpha, \beta)$ are $\hat{\alpha} = \frac{\bar{X}^2}{\left(\frac{1}{n}\right)\sum_{i=1}^{n}(X_i - \bar{X})^2}$ and $\hat{\beta} = \frac{\left(\frac{1}{n}\right)\sum_{i=1}^{n}(X_i - \bar{X})^2}{\bar{X}}$

e. Suppose that $X_1, \dots, X_n$ are iid with common pdf $f(x) = \begin{cases}(\theta+1)x^\theta, & 0 < x < 1 \\ 0 & , o.w.\end{cases}$ where $\theta (> 0)$ is the unknown parameter. Derive an estimator for θ by the method of moments.

5. Maximum Likelihood Estimation
    a. Likelihood function $\left(L(\vec{\theta})\right)$
    b. If $X \sim Bin(n, \theta)$ with n known, find $\hat{\theta}_{MLE}$.
    c. Suppose $X_1, \dots, X_n$ are iid $Bin(n, \theta)$ with n known, find $\hat{\theta}_{MLE}$
    d. Suppose $X_1, \dots, X_n$ are iid $N(\mu, \sigma^2)$, show that $\hat{\mu}_{MLE} = \bar{x}$ and $\hat{\sigma}^2{}_{MLE} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$
6. The Invariance Principle for MLE's $\left(\widehat{g(\theta)_{MLE}}) = g\left(\hat{\theta}_{MLE}\right)\right)$ (the function doesn't even have to be one-to-one!)
    a. From Number 5, part (c), find the MLE for $\theta^2 - 5\theta + 2$
    b. From Number 5, part (d), find the MLE for σ

Lecture – Problem Session for Ch 6 (10/26/18)

Lecture – Ch 7 Lecture 1 Slides 1 – 7 (10/29/18)

1. MLE Review
2. General Form Of Confidence Intervals (point est ± (critical value)(s.e. of point estimate))
3. Confidence Intervals vs Prediction Intervals
4. Confidence Interval for Mean of Normal Population $\left(\bar{x} \pm z_{\frac{\alpha}{2}}\left(\frac{\sigma}{\sqrt{n}}\right)\right)$
    a. Interpretation
    b. Factors impacting the width

Lecture – Halloween Jeopardy (10/31/18)

Lecture – Ch 7 Lecture 2 Slides 1 – 19 (11/2/18)

1. Confidence Intervals vs Prediction Intervals
2. General Form Of Confidence Intervals (point est ± (critical value)(s.e. of point estimate))
3. Confidence Interval for Mean μ of Normal Population $\left(\bar{x} \pm z_{\frac{\alpha}{2}}\left(\frac{\sigma}{\sqrt{n}}\right)\right)$
    a. Interpretation

4. Margin of Error & Factors impacting this
5. Minimum Sample Size $\left(n = \left(z_{\frac{\alpha}{2}}\frac{\sigma}{m}\right)^2\right)$ (Caution: your book defines width, not m)
6. Large sample (n > 40) C.I. for the Mean μ (regardless of the distribution of the original distribution) $\left(\bar{x} \pm z_{\frac{\alpha}{2}}\left(\frac{s}{\sqrt{n}}\right)\right)$
7. Confidence Interval for Population Proportion (for large n, i.e. $n\hat{p} \geq 10$ and $n(1-\hat{p}) \geq 10$) $\left(\hat{p} \pm z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right)$
8. Example 1 – Consider a normal population distribution with the value of σ know.
    a. What is the confidence level for the interval $\bar{x} \pm 2.81\frac{\sigma}{\sqrt{n}}$?
    b. What is the confidence level for the interval $\bar{x} \pm 1.44\frac{\sigma}{\sqrt{n}}$?
    c. What value of $z_{\frac{\alpha}{2}}$ in the interval $\bar{x} \pm z_{\frac{\alpha}{2}}\left(\frac{\sigma}{\sqrt{n}}\right)$ results in a confidence level of 99.7%? (hint: $1 - \frac{\alpha}{2} = \frac{1+CC}{2}$)
    d. What value of $z_{\frac{\alpha}{2}}$ in the interval $\bar{x} \pm z_{\frac{\alpha}{2}}\left(\frac{\sigma}{\sqrt{n}}\right)$ results in a confidence level of 75%?
9. Example 2 – On the basis of extensive tests, the yield point of a particular type of mild steel-reinforcing bar is known to be normally distributed with $\sigma = 100$. The composition of bars has been slightly modified, but the modification is not believed to have affected either the normality or the value of σ.
    a. Assuming this to be the case, if a sample of 25 modified bars resulted in a sample average yield point of 8439 lb, compute a 90% CI for the true average yield point of the modified bar.
    b. How would you modify the interval in part (a) to obtain a confidence level of 92%?
10. Determine the confidence level for each of the following large sample one-sided confidence bounds:
    a. Upper bound: $\bar{x} + (.84)\left(\frac{s}{\sqrt{n}}\right)$
    b. Lower bound: $\bar{x} - (2.05)\left(\frac{s}{\sqrt{n}}\right)$
    c. Upper bound: $\bar{x} + (.67)\left(\frac{s}{\sqrt{n}}\right)$

Lecture – Ch 7 Lecture 3 Slides 20 – 37 (11/5/18)

1. Determine the confidence level for each of the following large sample one-sided confidence bounds:
    a. Upper bound: $\bar{x} + (.84)\left(\frac{s}{\sqrt{n}}\right)$
    b. Lower bound: $\bar{x} - (2.05)\left(\frac{s}{\sqrt{n}}\right)$
    c. Upper bound: $\bar{x} + (.67)\left(\frac{s}{\sqrt{n}}\right)$
2. 1-sided confidence intervals

3. What if we don't know sigma and we have a small sample size (n < 40)? $\left(\bar{x} \pm t_{\frac{\alpha}{2},n-1}\frac{s}{\sqrt{n}}\right)$
4. Example – Determine the t critical value for the two-sided confidence interval in each of the following situations:
    a. Confidence level = 95%, df. = 10
    b. Confidence level = 99%, df. = 15
    c. Confidence level = 99%, df. = 38
5. Prediction Interval for future observation X_{n+1} – data from a normal distribution with known variance $\left(\bar{x} \pm z_{\frac{\alpha}{2}}\sigma\sqrt{1+\frac{1}{n}}\right)$
6. Prediction Interval for future observation X_{n+1} – data from a normal distribution with unknown variance $\left(\bar{x} \pm t_{\frac{\alpha}{2},n-1}s\sqrt{1+\frac{1}{n}}\right)$
7. Confidence Interval for σ^2 from a normal distribution with unknown variance $\left(\frac{vs^2}{\chi^2_{\frac{\alpha}{2},v}}, \frac{vs^2}{\chi^2_{1-\frac{\alpha}{2},v}}\right)$ (for σ, just take the square root of the interval).
8. Determine the following
    a. The 95th percentile of the Chi-Squared Distribution with 10 d.f.
    b. The P(10.98 ≤ χ^2 ≤ 36.78) where χ^2 has 22 d.f.

Lecture – Problem Session for Ch 7 (11/7/18)

Lecture – Quiz #3 (11/9/18)

Lecture – Quiz #3 Review (11/12/18)

Lecture – Ch 8 Lecture 1 - Slides 1 – 15 (11/14/18)

1. Terminology
    a. Statistical Hypothesis → A claim about a population, whether it is about a single parameter, the values of several parameters, or the form of an entire probability distribution
    b. Hypothesis Test → An assessment of the evidence provided by a data set in favor of (or against) a hypothesis about a population
    c. Test Statistic → the sample statistic. We want to see if the sample statistic is consistent with a hypothesis on the corresponding population parameter. (the entire goal is to see if this difference is statistically significant at the alpha significance level).
2. General Form of every Hypothesis test
3. The flow of every hypothesis test → Think of it as a proof by contradiction.
4. Two-sided vs one-sided alternatives

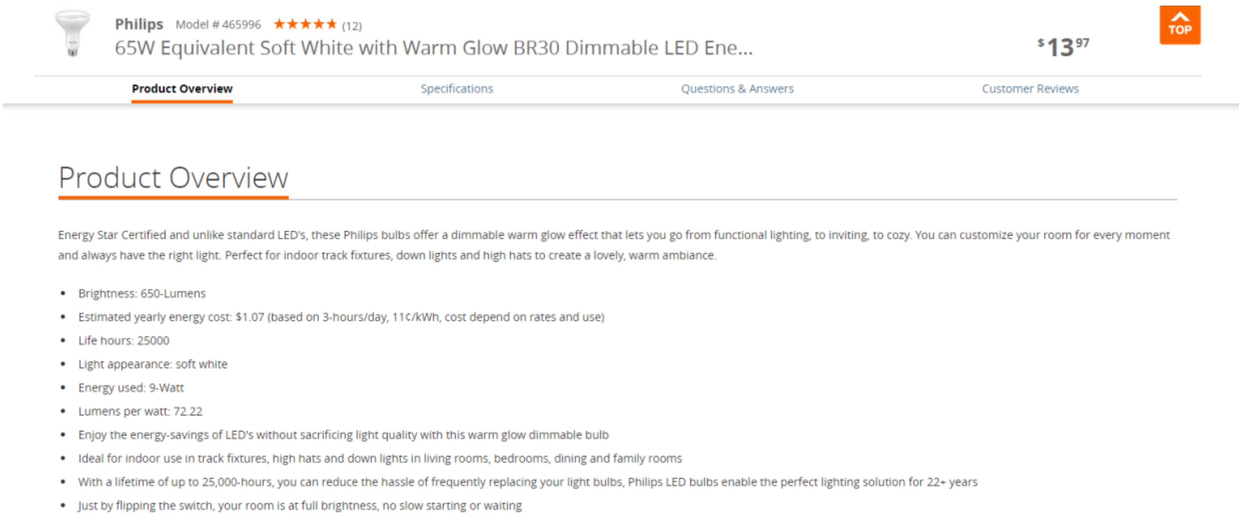
5. Example 1 – New Orleans is sinking. A study by the ASCE Journal of Hydrologic Engineering in 2016 cited that about 65% of New Orleans proper is at or below sea level. Hurricane Katrina highlighted the importance of levees in protecting the city. However, Professor Raymond Seed of the University of California, Berkeley, claims that during Katrina a surge of water estimated at 24 feet (about 10 feet higher than the levees along the city's eastern flank), swept into New Orleans from the Gulf of Mexico, causing most of the flooding in the city. The city now has put forth a policy that levees must be at least 10 feet higher than resting water level. How can we study this problem?
6. Types of Error
   a. Alpha (α) = P(type I error) = P(reject H_0 | H_0 is true)
   b. Beta (β) = P(type II error) = P(don't reject H_0 | H_1 is true)
   c. Power of a test = $1 - \beta$
   d. The goal: minimize alpha and beta. For a fixed alpha, we want the hypothesis test with the smallest beta.
   e. Trade-off between α and β → For fixed experiment (and sample size), decreasing alpha will increase beta, and vice versa.
7. Example 1 Cont. – What is a type 1 error in this case? What is a type 2 error in this case? Given the natural trade-off between alpha and beta, how should you choose to weight these?
8. Example 2 – Consider a population with the pdf $N(\theta, 1)$ where $\theta \epsilon \mathbb{R}$ is unknown. An experimenter wishes to test $H_0: \theta = 5.5$ vs $H_1: \theta = 8$ by collecting a random sample of $\vec{X} = (X_1, X_2, \dots, X_9)$ and is debating which test to use of the following:
   i. Reject H_0 iff $X_1 > 7$
   ii. Reject H_0 iff $\frac{1}{2}(X_1 + X_2) > 7$
   iii. Reject H_0 iff $\bar{X} > 6$
   a. Calculate alpha and beta for tests 1,2, and 3.
   b. Which test should you use, test 1 or test 3?
   c. Which test should you use, test 1 or test 2?
   d. What is the power of tests 1 and 3? What does this tell you?
   e. What if I change the distribution to a Chi squared distribution with degrees of freedom θ? How about a Binomial with $p = \theta$ and n known? How would the calculation change?

Lecture – Ch 8 Lecture 2 - Slides 16 – 26 (11/16/18)

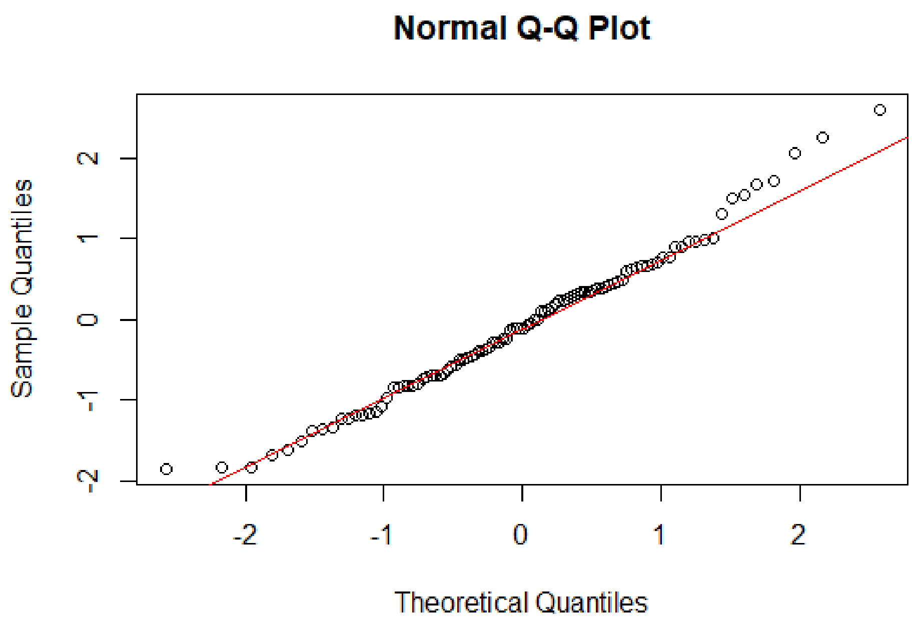
1. 1-Sample tests about a population mean
   a. 1-sample z-test → Normal distribution with known σ^2
   b. Large-Sample Approximate 1-Sample Z-Test → n > 40 unknown σ^2
   c. 1-Sample T-Test → unknown σ^2 (small n)
2. Summary

	1-Sample Z-Test	Large Sample Approx 1-Sample Z-test	1-Sample T-Test
Test Stat	$z^* = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$	$z^* = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$	$t^* = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$
$H_1: \mu > \mu_0$	$RR: z^* \geq z_\alpha$	$RR: z^* \geq z_\alpha$	$RR: t^* \geq t_{\alpha,n-1}$
$H_1: \mu < \mu_0$	$RR: z^* \leq -z_\alpha$	$RR: z^* \leq -z_\alpha$	$RR: t^* \leq -t_{\alpha,n-1}$
$H_1: \mu \neq \mu_0$	$RR: \lvert z^* \rvert \geq z_{\frac{\alpha}{2}}$	$RR: \lvert z^* \rvert \geq z_{\frac{\alpha}{2}}$	$RR: \lvert t^* \rvert \geq t_{\frac{\alpha}{2},n-1}$

3. The Process:
   a. Confirm Assumptions:
      i. One sample compared to known value
      ii. Testing for true unknown mean (μ)
      iii. In this case we have unknown σ^2
      iv. Do we have approximate normality?? (Normal Q-Q Plot)
4. Example 1 – The statistics department is ordering markers for the upcoming semester and needs to know the average lifetime of a marker for our instructors. The average lifetime posted on the company's website is 3 weeks, give or take 2 days assuming normality. However, the average lifetime for Austin's 25 markers last semester was 19 days, give or take 1 day. Can the department trust the company's website? Test your claim at the alpha significance level of .10.
5. Example 2 – In an effort to study the population of the nearly extinct mountain gorilla, researchers looked at the lifespan of these gorillas. Researchers studied 100 of the 880 remaining mountain gorillas and recorded their average age of 13 years old, give or take 2 years. Last year, the average age of mountain gorillas was 13.4 years. Is there reason to be concerned about the populations reproduction at the alpha significance level of .05?
6. Example 3 (Quality Assurance) - Philips produces 65W Dimmable LED Energy Star Light Bulbs sold at Home Depot. On the Home Depot site, they advertise the "life hours" of each light bulb is 25000:



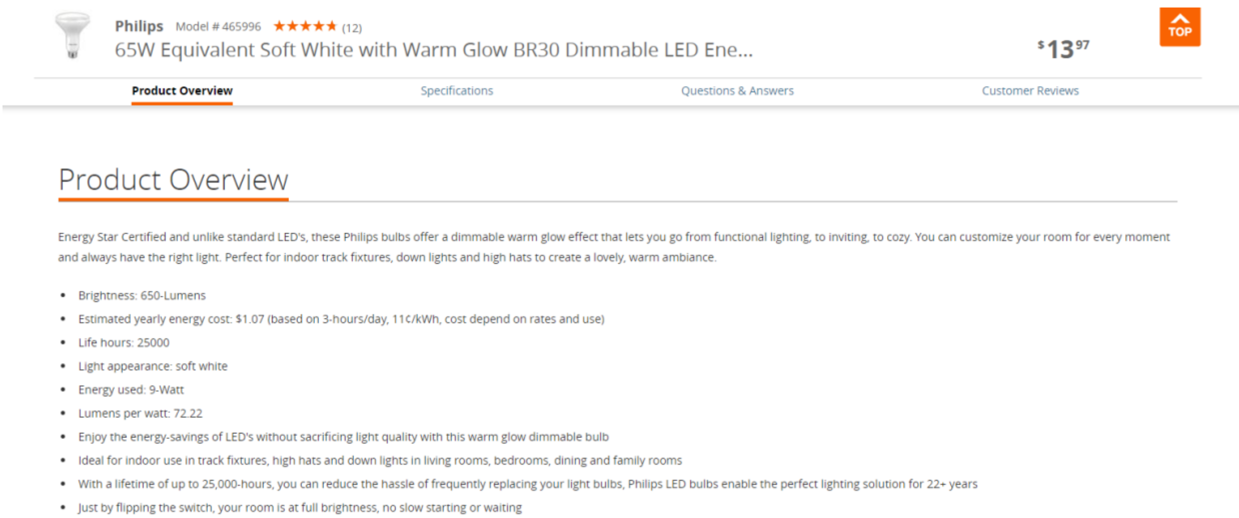
Accounting for variability, is the mean lifetime of light bulbs actually 25000? Assume we have a sample of n=100 light bulbs with $\bar{x} = 23024$ and sample st dev (s) = 6705. Also assume we have the following output:



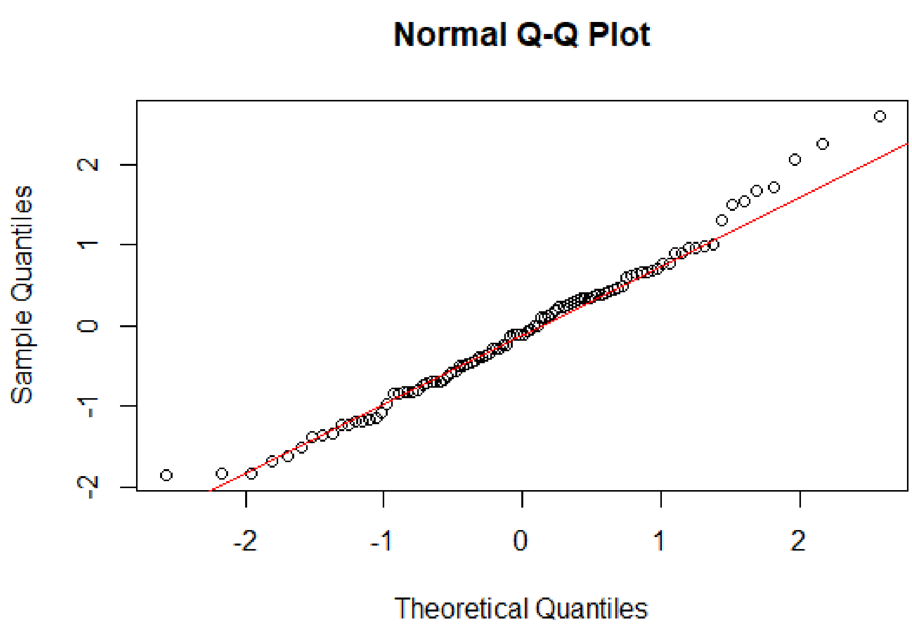
7. What happens when we don't have approximate normality? (Wilcoxon Signed Rank Test → Nonparametric (distribution free)
    a. Hypotheses test median not mean
    b. Test stat V = sum of positive signed ranks
    c. Rejection Rule → uses normal distribution

Lecture – Ch 8 Lecture 3 - Slides 27 – 41 (11/26/18)

1. SET Surveys
2. Example 3 (Quality Assurance) - Philips produces 65W Dimmable LED Energy Star Light Bulbs sold at Home Depot. On the Home Depot site, they advertise the "life hours" of each light bulb is 25000:



Accounting for variability, is the mean lifetime of light bulbs actually 25000? Assume we have a sample of n=100 light bulbs with $\bar{x} = 23024$ and sample st dev (s) = 6705. Also assume we have the following output:



3. 1-Sample Test for Population Proportion – Large-Sample Approximate Z-Test
    a. Use if $np_0 \geq 10$ and $n(1-p_0) \geq 10$
    b. (Don't worry too much about small sample test for proportion)
    c. Summary

	1-Sample Z-Test for Population Proportion
Test Stat	$z^* = \dfrac{\hat{p}-p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$
$H_1: p > p_0$	$RR: z^* \geq z_\alpha$
$H_1: p < p_0$	$RR: z^* \leq -z_\alpha$
$H_1: p \neq p_0$	$RR: \lvert z^* \rvert \geq z_{\frac{\alpha}{2}}$

4. Example 4 (Text Ch 8, Question 43) – A plan for an executive travelers' club has been developed by an airline on the premise that that 5% of its current customers would qualify for membership. A random sample of 500 customers yielded 40 who would qualify. Using this data, test at the alpha significance level of .01 the null hypothesis that the company's premise is correct against the alternative that it is not correct.
5. A Note on p-values and multiple hypothesis testing (Bonferroni Correction)
    a. **Quick Source:** https://www.npr.org/sections/thetwo-way/2016/09/13/493739074/50-years-ago-sugar-industry-quietly-paid-scientists-to-point-blame-at-fat
    b. **Further Reading:** Huff, Darrell, and Irving Geis (illustrator). *How to Lie with Statistics*. W.W. Norton & Co., 2006
        i. Discusses Bias, "A Well Chosen Average", Deceptive Visualizations, etc.

Lecture – Problem Session for Ch 8 (11/28/18)

Lecture – Ch 9 Lecture 1 - Slides 1 – 9 (11/30/18)

1. 2-Sample Z-Test with C.I. – comparing 2 population means (LABEL EACH)
    a. Assumptions
        i. 2 independent samples
        ii. Comparing means
        iii. Approximate normality for both samples (of both sampling distributions of sample means)

iv. Both with known, unequal, variances

b. Summary

	2-Sample Z-Test with C.I.
Test Stat	$z^* = \dfrac{\bar{x} - \bar{y} - d_0}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}}$
$H_1: \mu_1 - \mu_2 > d_0$	$RR: z^* \geq z_\alpha$
$H_1: \mu_1 - \mu_2 < d_0$	$RR: z^* \leq -z_\alpha$
$H_1: \mu_1 - \mu_2 \neq d_0$	$RR: \lvert z^* \rvert \geq z_{\frac{\alpha}{2}}$

c. Constructing the C.I. $\left((\bar{x} - \bar{y}) \pm z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}\right)$ for $\mu_1 - \mu_2$

i. Can think of this as (p.e. ±(crit value)(s.e. estimator))

2. Large-Sample Approx. 2-Sample Z-Test with C.I. – comparing 2 population means (LABEL EACH)

a. Assumptions

i. 2 independent samples

ii. Comparing means

iii. Approximate normality for both samples (of both sampling distributions of sample means)

iv. Both with unknown, unequal, variances (n,m ≥ 40)

b. Summary

	Large Sample Approx. 2-Sample Z-Test with C.I.
Test Stat	$z^* = \dfrac{\bar{x} - \bar{y} - d_0}{\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}}$
$H_1: \mu_1 - \mu_2 > d_0$	$RR: z^* \geq z_\alpha$
$H_1: \mu_1 - \mu_2 < d_0$	$RR: z^* \leq -z_\alpha$
$H_1: \mu_1 - \mu_2 \neq d_0$	$RR: \lvert z^* \rvert \geq z_{\frac{\alpha}{2}}$

c. Constructing the C.I. $\left((\bar{x}-\bar{y}) \pm z_{\frac{\alpha}{2}}\sqrt{\frac{s_1^2}{m}+\frac{s_2^2}{n}}\right)$ for $\mu_1 - \mu_2$
    i. Can think of this as (p.e. ±(crit value)(s.e. estimator))

3. Example 1 – (Textbook Chapter 9, Exercise 10) – An experiment was performed to compare the fracture toughness of high-purity 18 Ni maraging steel with commercial-purity steel of the same type (*Corrosion Science*, 1971: 723 – 736). For 32 specimens, the sample average toughness was 65.6 for the high-purity steel, whereas for the 38 specimens of commercial steel 59.8. Because the high-purity steel is more expensive, its use for a certain application can be justified only if its fracture toughness exceeds that of commercial purity steel by more than 5. Suppose that both toughness distributions are normal with the standard deviation of the high-purity steel known to be 1.2 and the standard deviation of the commercial steel known to be 1.1. Test this at an alpha significance level of .001.
    a. What is the appropriate test to use here? Justify your claim by confirming the assumptions are met.
    b. What are the null and alternative hypotheses?
    c. Calculate the test stat and the reject region. What is your conclusion at the alpha significance level stated?
    d. Calculate the p-value. What is your conclusion at the alpha significance level stated?
    e. What is a Type I error in context in this case? What is a Type II error in context in this case?
4. Example 2 – Suppose that instead we knew only that the sample standard deviation of the high-purity steel was known to be 1.2 and the sample standard deviation of the commercial steel was known to be 1.1. Suppose also that instead of 32 high-purity steel specimens we had 41, and instead of 38 specimens of commercial steel we had 52. How would our test change?

Lecture – Ch 9 Lecture 2 - Slides 10 – 19 (12/3/18)

1. Example 2 – Suppose that instead we knew only that the sample standard deviation of the high-purity steel was known to be 1.2 and the sample standard deviation of the commercial steel was known to be 1.1. Suppose also that instead of 32 high-purity steel specimens we had 41, and instead of 38 specimens of commercial steel we had 52. How would our test change?
2. Two-Sample t-test - comparing 2 population means (LABEL EACH)
    a. Assumptions
        i. 2 independent samples
        ii. Comparing means
        iii. Approximate normality for both samples (of both sampling distributions of sample means)
        iv. Both with unknown, unequal, variances (small m & n (<40))
    b. Summary

	2-Sample t-Test with C.I.
Test Stat	$t^* = \dfrac{\bar{x} - \bar{y} - d_0}{\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}}$ With d.f. $v = \dfrac{\left(\frac{s_1^2}{m} + \frac{s_2^2}{n}\right)^2}{\frac{\left(\frac{s_1^2}{m}\right)^2}{m-1} + \frac{\left(\frac{s_2^2}{n}\right)^2}{n-1}}$ (always round v down, regardless of decimal)
$H_1: \mu_1 - \mu_2 > d_0$	$RR: t^* \geq t_{\alpha,v}$
$H_1: \mu_1 - \mu_2 < d_0$	$RR: t^* \leq -t_{\alpha,v}$
$H_1: \mu_1 - \mu_2 \neq d_0$	$RR: \lvert t^* \rvert \geq t_{\frac{\alpha}{2},v}$

c. Constructing the C.I. $\left((\bar{x} - \bar{y}) \pm t_{\frac{\alpha}{2},v}\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}\right)$ for $\mu_1 - \mu_2$
    i. Can think of this as (p.e. ±(crit value)(s.e. estimator))

3. Pooled t-test - comparing 2 population means (LABEL EACH)
    a. Assumptions
        i. 2 independent samples
        ii. Comparing means
        iii. Approximate normality for both samples (of both sampling distributions of sample means)
        iv. Both with unknown variances (small m & n), but known that $\sigma_1^2 = \sigma_2^2 = \sigma^2$ (equal variance).
    b. Summary

	Pooled t-Test with C.I.
Test Stat	$t^* = \dfrac{\bar{x} - \bar{y} - d_0}{s_p\sqrt{\frac{1}{m} + \frac{1}{n}}}$ With $s_p^2 = \frac{m-1}{m+n-2}s_1^2 + \frac{n-1}{m+n-2}s_2^2$
$H_1: \mu_1 - \mu_2 > d_0$	$RR: t^* \geq t_{\alpha,m+n-2}$
$H_1: \mu_1 - \mu_2 < d_0$	$RR: t^* \leq -t_{\alpha,m+n-2}$

$H_1: \mu_1 - \mu_2 \neq d_0$	$RR: \lvert t^* \rvert \geq t_{\frac{\alpha}{2}, m+n-2}$

c. Constructing the C.I. $\left((\bar{x} - \bar{y}) \pm t_{\frac{\alpha}{2}, m+n-2} s_p \sqrt{\frac{1}{m} + \frac{1}{n}} \right)$ for $\mu_1 - \mu_2$

i. Can think of this as (p.e. ±(crit value)(s.e. estimator))

4. Example 3 (Textbook Ch 9, Exercise 25) – Low-back pain (LBP) is a serious health problem in many industrial settings. The article "Isodynamic Evaluation of Trunk Muscles and Low-Back Pain Among Workers in a Steel Factory" (*Ergonomics*, 1995: 2107-2117) reported the accompanying summary data on lateral range of motion (in degrees) for a sample without a history of LBP and another sample with a history of this malady:

Condition	Sample Size	Sample Mean	Sample SD
No LBP	28	91.5	5.5
LBP	31	88.3	7.8

a. What is the appropriate test to use here if we want to test whether lateral motion differs for the two conditions? Justify your claim by confirming the assumptions are met.
b. What are the null and alternative hypotheses?
c. Calculate the test stat and the reject region. What is your conclusion at the alpha significance level of .10?
d. Calculate the p-value. What is your conclusion at the alpha significance level of .15?
e. What is a Type I error in context in this case? What is a Type II error in context in this case?
f. Calculate a 90% confidence interval for the difference between population mean extent of lateral motion for the two conditions. Does the interval suggest that population mean lateral motion differs for the two conditions? Is the message different if a confidence level of 95% is used?
g. How would this example change if for some reason we assumed that the 2 populations had the same variance?

Lecture – Ch 9 Lecture 3 - Slides 20 – 29 (12/5/18)

1. Independent Samples vs Matched Pairs (linking factor)
2. Example 1 – Determine whether the samples in the following examples are independent or matched pairs:
   a. Researchers are interested in testing the effect of drinking alcohol on driving. So, they measure the time taken for 30 subjects each to complete a driving course. Then, each subject drinks 2 beers in 5 minutes and the researcher times the driving course again.

b. Researchers are interested in testing the effect of drinking alcohol on driving. So, they randomly select 30 U.S. males to complete a timed driving course. Then, they randomly select another 30 males to each drink 2 beers in 5 minutes and take the driving course.
c. Researchers are interested in testing the effect of drinking alcohol on driving. So, they randomly select 30 females and measure the time taken for each to complete a driving course. The goal of the study is to see if female times are lower than the national average of 1.34 minutes.

3. Paired t-test - comparing pop means (LABEL HOW YOU TAKE THE DIFFERENCE μ_d)
   a. **Inference is based on the average of the differences, not the difference of the averages!!**
   b. Assumptions
      i. 2 samples ***NOT*** independent (have some linking factor)
      ii. Comparing means
      iii. The sample of differences comes from a normal population (CLT: if n ≥ 30 then always safe to use the test regardless of the shape of the distribution. If n < 30, then can check the shape of the distribution with a histogram to ensure that it's approx. normal)
   c. Summary

	Paired t-Test with C.I.
Test Stat	$t^* = \dfrac{\bar{x}_d - d_0}{\frac{s_d}{\sqrt{n}}}$ With $d.f. = n-1$
$H_1: \mu_d > d_0$	$RR: t^* \geq t_{\alpha, n-1}$
$H_1: \mu_d < d_0$	$RR: t^* \leq -t_{\alpha, n-1}$
$H_1: \mu_d \neq d_0$	$RR: \lvert t^* \rvert \geq t_{\frac{\alpha}{2}, n-1}$

   d. Constructing the C.I. $\left(\bar{x}_d \pm t_{\frac{\alpha}{2}, n-1} \frac{s_d}{\sqrt{n}}\right)$ for μ_d
      i. Can think of this as (p.e. ±(crit value)(s.e. estimator))

4. 2 Sample Proportion Z-test - comparing pop proportions (LABEL EACH)
   a. Assumptions
      i. 2 independent samples
      ii. Comparing proportions
      iii. $m\widehat{p_1}, m(1-\widehat{p_1}), n\widehat{p_2}, n(1-\widehat{p_2}) \geq 10$
   b. Summary

	2-Sample Proportion Z-test with C.I.

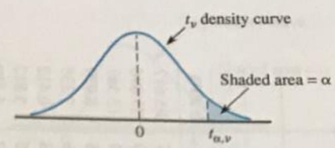
Test Stat	$z^* = \dfrac{\widehat{p_1} - \widehat{p_2} - p_0}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{m} + \frac{1}{n}\right)}}$ with $\hat{p} = \frac{m}{m+n}\widehat{p_1} + \frac{n}{m+n}\widehat{p_2}$
$H_1: p_1 - p_2 > p_0$	$RR: z^* \geq z_\alpha$
$H_1: p_1 - p_2 < p_0$	$RR: z^* \leq -z_\alpha$
$H_1: p_1 - p_2 \neq p_0$	$RR: \lvert z^* \rvert \geq z_{\frac{\alpha}{2}}$

c. Constructing the C.I. $\left((\widehat{p_1} - \widehat{p_2}) \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\widehat{p_1}(1-\widehat{p_1})}{m} + \frac{\widehat{p_2}(1-\widehat{p_2})}{n}} \right)$ for $p_1 - p_2$

i. Can think of this as (p.e. ±(crit value)(s.e. estimator))

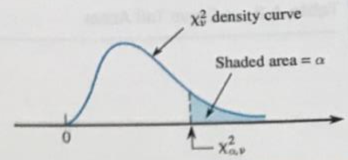
5. Hypothesis Testing Flowchart – When to use each test

Table A.5 Critical Values for *t* Distributions



ν \ α	.10	.05	.025	.01	.005	.001	.0005
1	3.078	6.314	12.706	31.821	63.657	318.31	636.62
2	1.886	2.920	4.303	6.965	9.925	22.326	31.598
3	1.638	2.353	3.182	4.541	5.841	10.213	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	1.319	1.714	2.069	2.500	2.807	3.485	3.767
24	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	1.310	1.697	2.042	2.457	2.750	3.385	3.646
32	1.309	1.694	2.037	2.449	2.738	3.365	3.622
34	1.307	1.691	2.032	2.441	2.728	3.348	3.601
36	1.306	1.688	2.028	2.434	2.719	3.333	3.582
38	1.304	1.686	2.024	2.429	2.712	3.319	3.566
40	1.303	1.684	2.021	2.423	2.704	3.307	3.551
50	1.299	1.676	2.009	2.403	2.678	3.262	3.496
60	1.296	1.671	2.000	2.390	2.660	3.232	3.460
120	1.289	1.658	1.980	2.358	2.617	3.160	3.373
∞	1.282	1.645	1.960	2.326	2.576	3.090	3.291

Table A.7 Critical Values for Chi-Squared Distributions

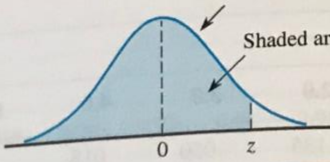


ν \ α	.995	.99	.975	.95	.90	.10	.05	.025	.01	.005
1	0.000	0.000	0.001	0.004	0.016	2.706	3.843	5.025	6.637	7.882
2	0.010	0.020	0.051	0.103	0.211	4.605	5.992	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.344	12.837
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.832	15.085	16.748
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.440	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.012	18.474	20.276
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.534	20.090	21.954
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.022	21.665	23.587
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.724	26.755
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	7.041	19.812	22.362	24.735	27.687	29.817
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4.600	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.577	32.799
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	5.697	6.407	7.564	8.682	10.085	24.769	27.587	30.190	33.408	35.716
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	6.843	7.632	8.906	10.117	11.651	27.203	30.143	32.852	36.190	38.580
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997
21	8.033	8.897	10.283	11.591	13.240	29.615	32.670	35.478	38.930	41.399
22	8.643	9.542	10.982	12.338	14.042	30.813	33.924	36.781	40.289	42.796
23	9.260	10.195	11.688	13.090	14.848	32.007	35.172	38.075	41.637	44.179
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.558
25	10.519	11.523	13.120	14.611	16.473	34.381	37.652	40.646	44.313	46.925
26	11.160	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290
27	11.807	12.878	14.573	16.151	18.114	36.741	40.113	43.194	46.962	49.642
28	12.461	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.993
29	13.120	14.256	16.147	17.708	19.768	39.087	42.557	45.772	49.586	52.333
30	13.787	14.954	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672
31	14.457	15.655	17.538	19.280	21.433	41.422	44.985	48.231	52.190	55.000
32	15.134	16.362	18.291	20.072	22.271	42.585	46.194	49.480	53.486	56.328
33	15.814	17.073	19.046	20.866	23.110	43.745	47.400	50.724	54.774	57.646
34	16.501	17.789	19.806	21.664	23.952	44.903	48.602	51.966	56.061	58.964
35	17.191	18.508	20.569	22.465	24.796	46.059	49.802	53.203	57.340	60.272
36	17.887	19.233	21.336	23.269	25.643	47.212	50.998	54.437	58.619	61.581
37	18.584	19.960	22.105	24.075	26.492	48.363	52.192	55.667	59.891	62.880
38	19.289	20.691	22.878	24.884	27.343	49.513	53.384	56.896	61.162	64.181
39	19.994	21.425	23.654	25.695	28.196	50.660	54.572	58.119	62.426	65.473
40	20.706	22.164	24.433	26.509	29.050	51.805	55.758	59.342	63.691	66.766

For $\nu > 40$, $\chi^2_{\alpha,\nu} \approx \nu\left(1 - \frac{2}{9\nu} + z_\alpha\sqrt{\frac{2}{9\nu}}\right)^3$

Table A.3 Standard Normal Curve Areas

$\Phi(z) = P(Z \leq z)$



z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0017	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0038
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0352	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0722	.0708	.0694	.0681
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
−1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
−1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
−0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
−0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
−0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
−0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
−0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
−0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
−0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3482
−0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
−0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
−0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

(continued)

Table A.3 Standard Normal Curve Areas *(cont.)*

$\Phi(z) = P(Z \leq z)$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9278	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998