INTERVAL ESTIMATION AND HYPOTHESES TESTING

1. BRIEF REVIEW OF SOME COMMON DISTRIBUTIONS

1.1. **Probability Distributions.** Table 1 contains information on some common probability distributions.

TABLE 1. Probability Distributions

Distribution	pdf	Mean	Variance	mgf	Notes
Normal, N(μ, σ^2)	$\frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$	μ	σ^2	$e^{\left(\mu t + \frac{t^2 \sigma^2}{2}\right)}$	$\mu = 0, \sigma^2 = 1$ $f(x;\mu,\sigma^2)=\frac{1}{\sqrt{2\pi}} \cdot e^{\frac{-1}{2}x^2}$ $M_X(t) = e^{\frac{t^2}{2}}$
Exponential	$\frac{1}{\theta} \cdot e^{\frac{-x}{\theta}}, \quad \theta > 0$	θ	θ^2	$(1-\theta t)^{-1}$ $\frac{1}{1-\theta t}$	
Chi-square, $\chi^2(\nu)$	$\frac{1}{2^{\frac{\nu}{2}} \Gamma\left(\frac{\nu}{2}\right)} x^{\frac{\nu}{2}-1} e^{\frac{-x}{2}}$	ν	2ν	$(1-2t)^{\frac{-\nu}{2}}$ $\left(\frac{1}{1-2t}\right)^{\frac{\nu}{2}}$	$X_i \sim N(0,1),\ i = 1, 2, \ldots, n$ $\Rightarrow \sum_{i=1}^{n} X_i^2 \sim \chi^2(n)$ $X_i \sim N(\mu, \sigma^2),\ i = 1, 2, \ldots, n$ $\Rightarrow \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma}\right)^2 \sim \chi^2(n)$ $\Rightarrow \sum_{i=1}^{n} \left(\frac{X_i - \bar{X}}{\sigma}\right)^2 \sim \chi^2(n-1)$
Gamma	$\frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{\frac{-x}{\beta}}$	$\alpha\beta$	$\alpha\beta^2$	$(1-\beta t)^{-\alpha}$ $\left(\frac{1}{1-\beta t}\right)^{\alpha}$	
t(ν)	$\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu}\,\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{x^2}{\nu}\right)^{\frac{-(\nu+1)}{2}}$	0	$\frac{\nu}{\nu-2}$		$t(\nu) = \frac{N(0,1)}{\sqrt{\frac{\chi^2(\nu)}{\nu}}}$ $N(0,1)$ and $\chi^2(\nu)$ are independent
F(ν_1,ν_2)	$\frac{\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)} \cdot \left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}} \cdot x^{\frac{\nu_1}{2}-1}$ $\times \left(1+\frac{\nu_1}{\nu_2}x\right)^{\frac{-(\nu_1+\nu_2)}{2}}$	$\frac{\nu_2}{\nu_2-2}$	$\frac{2\nu_2^2(\nu_1+\nu_2-2)}{\nu_1(\nu_2-2)^2(\nu_2-4)}$		$F(\nu_1,\nu_2)=\frac{\frac{\chi_1^2(\nu_1)}{\nu_1}}{\frac{\chi_2^2(\nu_2)}{\nu_2}} = \frac{\nu_2}{\nu_1} \cdot \frac{\chi_1^2(\nu_1)}{\chi_2^2(\nu_2)}$ $\chi_1^2(\nu_1)$ and $\chi_1^2(\nu_2)$ are independent

1.2. **Moments.**

1.2.1. *Population raw moments.*

$$
\begin{aligned}
\mu_r' &= E(X^r) = \int_{-\infty}^{\infty} x^r f(x)\, dx \\
\mu_1' &= E(X) = \int_{-\infty}^{\infty} x f(x)\, dx
\end{aligned}
\tag{1}
$$

Date: February 18, 2008.

1.2.2. *Population central moments.*

$$\begin{aligned} \mu_r &= E[(X - \mu_X)^r] = \int_{-\infty}^{\infty} (x - \mu_X)^r f(x)\,dx \\ \mu_1 &= E[X - \mu_X] = \int_{-\infty}^{\infty} (x - \mu_X) f(x)\,dx \\ \mu_2 &= E[(X - \mu_X)^2] = \int_{-\infty}^{\infty} (x - \mu_X)^2 f(x)\,dx = Var(X) = \sigma^2 \end{aligned} \tag{2}$$

1.2.3. *Sample raw moments.*

$$\begin{aligned} \bar{X}_n^r &= \frac{1}{n}\sum_{i=1}^{n} X_i^r \\ \bar{X} &= \frac{1}{n}\sum_{i=1}^{n} X_i \end{aligned} \tag{3}$$

1.2.4. *Sample central moments.*

$$\begin{aligned} C_n^r &= \frac{1}{n}\sum_{i=1}^{n} \left(X_i - \mu'_{i,1}\right)^r \quad r = 1, 2, 3, \ldots, \\ C_n^1 &= \frac{1}{n}\sum_{i=1}^{n} \left(X_i - \mu'_{i,1}\right) \\ C_n^2 &= \frac{1}{n}\sum_{i=1}^{n} \left(X_i - \mu'_{i,1}\right)^2 \end{aligned} \tag{4}$$

1.2.5. *Sample moments about the average (sample mean).*

$$\begin{aligned} M_n^r &= \frac{1}{n}\sum_{i=1}^{n} \left(X_i - \bar{X}_n\right)^r, \quad r = 1, 2, 3, \ldots, \\ M_n &= \frac{1}{n}\sum_{i=1}^{n} \left(X_i - \bar{X}_n\right) = 0 \\ M_n^2 &= \frac{1}{n}\sum_{i=1}^{n} \left(X_i - \bar{X}_n\right)^2 = \hat{\sigma}^2 \end{aligned} \tag{5}$$

1.3. **Properties of various sample moments and functions thereof.** Consider a sample $X_1, X_2, \ldots, X_n$ where the X_i are identically and independently distributed with mean μ and variance σ^2.

Now consider a sample $X_1, X_2, \ldots, X_n$ where the X_i are identically and independently distributed normal random variables with mean μ and variance σ^2.

2. The Basic Idea of Interval Estimates

An interval rather than a point estimate is often of interest. Confidence intervals are thus important in empirical work. To construct interval estimates, standardized normal random variables are often used.

TABLE 2. Properties of sample moments

Moment	**Expected Value**	**Variance**
$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$	μ	$\frac{\sigma^2}{n}$
$C_n^1 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)$	0	$\frac{\sigma^2}{n}$
$C_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2$	σ^2	$\frac{1}{n}Var\left[(X-\mu)^2\right] = \frac{E(X-\mu)^4 - \left[E(X-\mu)^2\right]^2}{n} = \frac{\mu_4 - \sigma^4}{n}$
$M_n^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2$	$\frac{n-1}{n}\sigma^2$	$\frac{(n-1)^2\mu_4}{n^3} - \frac{(n-1)(n-3)\sigma^4}{n^3} = \frac{\mu_4-\mu_2^2}{n} - \frac{2(\mu_4-2\mu_2^2)}{n^2} + \frac{\mu_4-3\mu_2^2}{n^3}$
$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2$	σ^2	$\frac{\mu_4}{n} - \frac{(n-3)\sigma^4}{n(n-1)}$

TABLE 3. Properties of sample moments of a normal distribution

Moment	**Value**	**Expected Value**	**Variance**
$\mu_1 = E(X-\mu)$	0	-	-
$\mu_2 = E(X-\mu)^2$	σ^2	-	-
$\mu_3 = E(X-\mu)^3$	0	-	-
$\mu_4 = E(X-\mu)^4$	$3\sigma^4$	-	-
$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$	-	μ	$\frac{\sigma^2}{n}$
$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2$	-	σ^2	$\frac{2\sigma^4}{n-1}$
$M_n^2 = \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2$	-	$\frac{n-1}{n}\sigma^2$	$\frac{2\sigma^4(n-1)}{n^2}$

3. STANDARDIZED NORMAL VARIABLES AND CONFIDENCE INTERVALS FOR THE MEAN WITH σ KNOWN

3.1. **Form of the confidence interval.** If X is a normal random variable with mean μ and variance σ^2 then

$$Z = \left(\frac{X-\mu}{\sigma}\right) \tag{6}$$

is a standard normal variable with mean zero and variance one. An estimate of μ, $\hat{\mu}$ is given by the sample mean $\bar{X}$. This then implies

$$\begin{aligned} \bar{X} &= \frac{1}{n}\sum_{i=1}^{n} X_i \;=\; \bar{X} \sim N\left(\mu,\, \frac{\sigma^2}{n}\right) \\ z &= \frac{\bar{X}-\mu}{\left[\frac{\sigma}{\sqrt{n}}\right]} \sim N(0,\,1) \end{aligned} \tag{7}$$

So if γ_1 is the upper $\alpha/2$ percent critical value of a standard normal variable, i.e.

$$\int_{\gamma_1}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}\, dz = \frac{\alpha}{2} \quad \text{then}$$

$$\begin{aligned} 1-\alpha = F(\gamma_1) - F(-\gamma_1) &= \Pr\left[-\gamma_1 \le \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}} \le \gamma_1\right] \\ &= \Pr\left[-\gamma_1 \frac{\sigma}{\sqrt{n}} \le \bar{x}-\mu \le \gamma_1 \frac{\sigma}{\sqrt{n}}\right] \\ &= \Pr\left[\gamma_1 \frac{\sigma}{\sqrt{n}} \ge -\bar{x}+\mu \ge -\gamma_1 \frac{\sigma}{\sqrt{n}}\right] \\ &= \Pr\left[\bar{x}-\gamma_1 \frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x}+\gamma_1 \frac{\sigma}{\sqrt{n}}\right] \end{aligned} \tag{8}$$

Therefore, with σ known,

$$\left[\bar{x}-\gamma_1 \frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x}+\gamma_1 \frac{\sigma}{\sqrt{n}}\right] \tag{9}$$

is said to be the $(1-\alpha)$ 100% confidence interval for μ.

3.2. **Example confidence interval for the mean of a normal distribution with σ^2 known.** Consider the income data for carpenters and house painters in table 4 where the data for both groups of individuals is distributed normally. Further assume that σ^2 is known and equal to $600,000.

TABLE 4. Income Data for Carpenters and House Painters

	carpenters	painters
sample size	n_c = 12	n_p = 15
mean income	$\bar{c}$ = $6000	$\bar{p}$ = $5400
estimated variance	s_c^2 = $565 000	s_p^2 = $362 500

Consider a 95% confidence interval for the mean of the distribution which we will denote by μ_c. For the normal distribution with $\frac{\alpha}{2}$ = 0.025, γ_1 is equal to 1.96. We then have

$$\begin{aligned} 1-\alpha &= \Pr\left[6000 - \frac{\sqrt{600\,000}}{\sqrt{12}}(1.96) \le \mu_c \le 6000 + \frac{\sqrt{600\,000}}{\sqrt{12}}(1.96)\right] \\ &= \Pr\left[6000 - (223.607)(1.96) \le \mu_x \le 6000 + (223.607)(1.96)\right] \\ &= \Pr\left[5561.73 \le \mu_x \le 6438.27\right] \end{aligned} \tag{10}$$

4. CONFIDENCE INTERVALS FOR THE MEAN WITH σ UNKNOWN

4.1. **Form of the confidence interval.** The previous section gave an interval estimate for the mean of a population when σ was known. When σ is unknown, another method must be used. Recall from the section 1 on probability distributions that the t random variable is defined as

$$t = \frac{z}{\sqrt{\frac{\chi^2(\nu)}{\nu}}} \tag{11}$$

where z is a standard normal and $\chi^2(\nu)$ is a χ^2 random variable with ν degrees of freedom and z and $\chi^2(\nu)$ are independent.

Also note from the same section 1 that if $X_i \sim N(\mu,\, \sigma^2)$ then

$$\begin{aligned} \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma} \right)^2 &\sim \chi^2(n) \\ \text{and} \quad \sum_{i=1}^{n} \left(\frac{X_i - \bar{X}}{\sigma} \right)^2 &\sim \chi^2(n-1) \end{aligned} \tag{12}$$

If $Y \sim N(\beta,\, \sigma^2)$ and $\bar{Y} = \hat{\beta}$, then we would have

$$\sum_{i=1}^{n} \left(\frac{Y_i - \hat{\beta}}{\sigma} \right)^2 \sim \chi^2(n-1) \tag{13}$$

We can use the information in equation 12 to find the distribution of $\frac{(n-1)S^2}{\sigma^2}$ where S^2 is equal to

$$S^2 \;=\; \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2 \tag{14}$$

Now substitute for S^2 in $\frac{(n-1)S^2}{\sigma^2}$ where $S^2$and simplify as follows.

$$\begin{aligned} \frac{(n-1)S^2}{\sigma^2} &= \frac{(n-1)}{\sigma^2} \cdot \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2 \\ &= \frac{1}{\sigma^2} \cdot \sum_{i=1}^{n} (X_i - \bar{X})^2 \\ &= \sum_{i=1}^{n} \left(\frac{X_i - \bar{X}}{\sigma} \right)^2 \sim \chi^2(n-1) \end{aligned} \tag{15}$$

The last line then indicates that

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1) \tag{16}$$

We can show that $\frac{\bar{X}-\mu}{\left[\frac{\sigma}{\sqrt{n}}\right]}$ in equation 7 and $\frac{(n-1)S^2}{\sigma^2}$ in equation 16 are independent so that

$$
\begin{aligned}
\frac{\frac{\frac{\bar{X}-\mu}{\sigma}}{\sqrt{n}}}{\sqrt{\frac{(n-1)S^2}{\sigma^2(n-1)}}} &= \frac{\frac{\frac{\bar{X}-\mu}{\sigma}}{\sqrt{n}}}{\frac{1}{\sigma}\sqrt{S^2}} \\
&= \frac{\bar{X}-\mu}{\frac{S}{\sqrt{n}}} = \frac{\sqrt{n}\,\left(\bar{X}-\mu\right)}{S} \sim t(n-1)
\end{aligned}
\tag{17}
$$

If γ_1 is the upper $\alpha/2$ percent critical value of a t random variable then

$$
\int_{\gamma_1}^{\infty} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu}\,\Gamma\left(\frac{\nu}{2}\right)} \left(1+\frac{t^2}{\nu}\right)^{-\frac{(\nu+1)}{2}} dt = \frac{\alpha}{2}
\tag{18}
$$

and

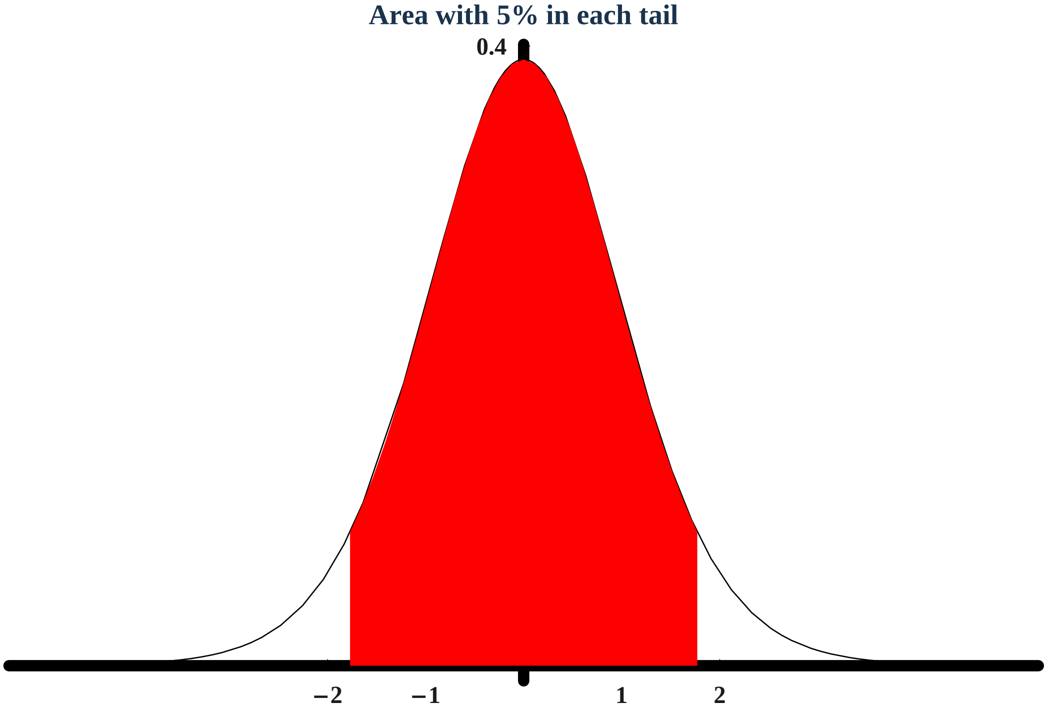
$$
\begin{aligned}
1-\alpha = F(\gamma_1) - F(-\gamma_1) &= \Pr\left[-\gamma_1 \leq \frac{\frac{\frac{\bar{X}-\mu}{\sigma}}{\sqrt{n}}}{\sqrt{\frac{(n-1)S^2}{\sigma^2(n-1)}}} \leq \gamma_1\right] \\
&= \Pr\left[-\gamma_1 \leq \frac{\sqrt{n}(\bar{X}-\mu)}{S} \leq \gamma_1\right] \\
&= \Pr\left[\frac{-\gamma_1 S}{\sqrt{n}} \leq \bar{X}-\mu \leq \frac{\gamma_1 S}{\sqrt{n}}\right] \\
&= \Pr\left[\bar{X} - \frac{\gamma_1 S}{\sqrt{n}} \leq \mu \leq \bar{X} + \frac{\gamma_1 S}{\sqrt{n}}\right]
\end{aligned}
\tag{19}
$$

This is referred to as $(1-\alpha)(100\%)$ confidence interval for μ. Figure 1 gives the area such that 5% of the distribution is the left of the shaded area and 5% of the distribution is the right of the shaded area.

4.2. **Example confidence interval for the mean of a normal distribution with S_c^2 = \$565 000.** Consider the income data for carpenters and house painters in table 4 where the data for both groups of individuals is distributed normally. We can construct the 95% confidence interval as follows:

$$
\begin{aligned}
1-\alpha &= \Pr\left[6000 - \frac{\sqrt{565\,000}}{\sqrt{12}}(2.201) \leq \mu_c \leq 6000 + \frac{\sqrt{565\,000}}{\sqrt{12}}(2.201)\right] \\
&= \Pr\left[6000 - (216.99)(2.201) \leq \mu_x \leq 6000 + (216.99)(2.201)\right] \\
&= \Pr\left[5522.4 \leq \mu_c \leq 6477.6\right]
\end{aligned}
\tag{20}
$$

FIGURE 1. Probability Density Function for a t Distribution



So, for example, 7000 is not in the 95% confidence interval for the mean.

5. CONFIDENCE INTERVALS FOR THE VARIANCE

5.1. **Form of the confidence interval.** Remember from equation 16 that

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1) \tag{21}$$

Now if γ_1 and γ_2 are such that

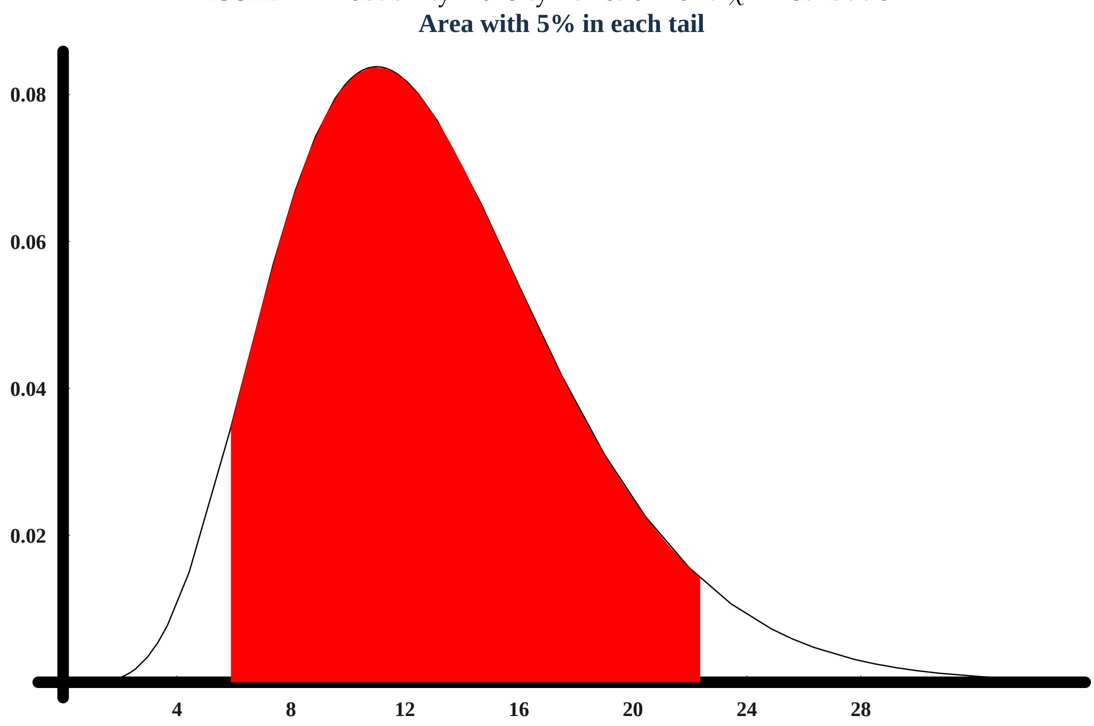
$$\begin{aligned} \Pr\big(\chi^2(\nu) \leq \gamma_1\big) &= \frac{\alpha}{2} \\ \Pr\big(\chi^2(\nu) \geq \gamma_2\big) &= \frac{\alpha}{2} \end{aligned} \tag{22}$$

Then

$$\begin{aligned} 1-\alpha = F(\gamma_2,\,\nu) - F(\gamma_1,\,\nu) &= \Pr\left[\gamma_1 \le \frac{(n-1)S^2}{\sigma^2} \le \gamma_2\right] \\ &= \Pr\left[\frac{\gamma_1\sigma^2}{n-1} \le S^2 \le \frac{\gamma_2\sigma^2}{n-1}\right] \\ &= \Pr\left[\frac{(n-1)}{\sigma^2}\gamma_1 \ge \frac{1}{S^2} \ge \frac{(n-1)}{\sigma^2\gamma_2}\right] \\ &= \Pr\left[\frac{(n-1)S^2}{\gamma_2} \le \sigma^2 \le \frac{(n-1)S^2}{\gamma_1}\right] \end{aligned} \tag{23}$$

is the $(1-\alpha)$ 100% confidence interval for the variance σ^2. Figure 2 gives the area such that 5% of the distribution is the left of the shaded area and 5% of the distribution is the right of the shaded area.

FIGURE 2. Probability Density Function for a χ^2 Distribution



5.2. **Example confidence interval for the variance a normal distribution.** Consider the income data for carpenters and house painters in table 4 where the data for both groups of individuals are distributed normally. We can construct the 95% confidence interval for the variance of the painters as follows:

$$1 - \alpha \; = \; \Pr\left[\frac{(n_p - 1)S_p^2}{\gamma_2} \leq \sigma_p^2 \leq \frac{(n_p - 1)S_p^2}{\gamma_1} \right] \tag{24}$$

For a χ^2 distribution with 14 degrees of freedom we obtain γ_1 = 5.63 and γ_2 = 26.12. We then obtain

$$\begin{aligned} 1 - \alpha \; &= \; \Pr\left[\frac{(14)(362\,500)}{26.12} \leq \sigma_p^2 \leq \frac{(14)(362\,500)}{5.63} \right] \\ &= \; \Pr\left[195\,042 \leq \sigma_p^2 \leq 901\,421 \right] \end{aligned} \tag{25}$$

6. Two Samples and a Confidence Interval for the Variances

6.1. **Form of the confidence interval.** Suppose that $x_1^1, x_2^1, \ldots, x_n^1$ is a random sample from a distribution $X_1 \sim N(\mu_1, \sigma_1^2)$ and $x_1^2, x_2^2, \ldots, x_n^2$ is a random sample from a distribution $X_2 \sim N(\mu_2, \sigma_2^2)$.

Now remember that the ratio of two chi-squared variables divided by their degrees of freedom is distributed as an F variable. For this example

$$\begin{aligned} \frac{(n_1 - 1)S_1^2}{\sigma_1^2} &\sim \chi^2(n_1 - 1) \\ \frac{(n_2 - 1)S_2^2}{\sigma_2^2} &\sim \chi^2(n_2 - 1) \end{aligned} \tag{26}$$

Now divide each of the chi-squared variables by its degrees of freedom and then take the ratio as follows

$$\frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} \sim F\left(n_1 - 1, \, n_2 - 1\right) \tag{27}$$

F distributions are normally tabled giving the area in the upper tail, i.e.

$$1 - \alpha = \Pr\left(F_{n_1 - 1, \, n_2 - 1} \leq \gamma\right) \quad \text{or} \quad \alpha = \Pr\left(F_{n_1 - 1, \, n_2 - 1} \geq \gamma\right) \tag{28}$$

Now let γ_1 and γ_2 be such that

$$\begin{aligned} \Pr\left(F_{n_1 - 1, \, n_2 - 1} \leq \gamma_1\right) &= \frac{\alpha}{2} \\ \Pr\left(F_{n_1 - 1, \, n_2 - 1} \geq \gamma_2\right) &= \frac{\alpha}{2} \\ \Pr\left(F_{n_1 - 1, \, n_2 - 1} \leq \gamma_2\right) &= 1 - \frac{\alpha}{2} \\ \Pr\left(F_{n_1 - 1, \, n_2 - 1} \leq \gamma_2\right) - \Pr\left(F_{n_1 - 1, \, n_2 - 1} \leq \gamma_1\right) &= 1 - \frac{\alpha}{2} - \frac{\alpha}{2} = 1 - \alpha \end{aligned} \tag{29}$$

Remembering that

$$F_{1 - \alpha/2, \, n_1 - 1, \, n_2 - 1} = \frac{1}{F_{\alpha/2, \, n_2 - 1, \, n_1 - 1}} \tag{30}$$

then if

$$\Pr\left(F_{n_1-1,\,n_2-1} \leq \gamma_2\right) = 1 - \frac{\alpha}{2}$$

then

$$\Pr(F_{n_2-1,\,n_1-1} \leq \frac{1}{\gamma_2}) = \frac{\alpha}{2} \tag{31}$$

We can now construct a confidence interval as follows:

$$\begin{aligned} 1-\alpha &= \Pr\left[\gamma_1 \leq \frac{S_1^2\sigma_2^2}{S_2^2\sigma_1^2} \leq \gamma_2\right] \\ &= \Pr\left[\frac{S_2^2}{S_1^2}\gamma_1 \leq \frac{\sigma_2^2}{\sigma_1^2} \leq \frac{S_2^2}{S_1^2}\gamma_2\right] \\ &= \Pr\left[\frac{S_1^2}{S_2^2}\frac{1}{\gamma_1} \geq \frac{\sigma_1^2}{\sigma_2^2} \geq \frac{S_1^2}{S_2^2}\frac{1}{\gamma_2}\right] \\ &= \Pr\left[\frac{S_1^2}{S_2^2}\frac{1}{\gamma_2} \leq \frac{\sigma_1^2}{\sigma_2^2} \leq \frac{S_1^2}{S_2^2}\frac{1}{\gamma_1}\right] \end{aligned} \tag{32}$$

This is then the $(1-\alpha)$ 100% confidence interval for ratio of the variances.

6.2. **Example confidence interval for the ratio of variances with α = .05.** We can use equation 32as follows

$$1-\alpha = \Pr\left[\frac{S_c^2}{S_p^2}\frac{1}{\gamma_2} \leq \frac{\sigma_c^2}{\sigma_p^2} \leq \frac{S_c^2}{S_p^2}\frac{1}{\gamma_1}\right] \tag{33}$$

The upper critical level is easy to obtain as $F(11,\, 14, : 0.025) = 3.10$ since we want the area to the left of the critical value to be .975. Since the tables don't contain the lower tail, we obtain the critical value $\gamma_1(11,\, 14)$ as $(1/\gamma_1(14,\, 11)$. This critical value is given by $F(14,\, 11\, :\, 0.025) = 3.36$. The reciprocal of this is .297. Notice that the confidence interval is in reciprocal form so we also need the reciprocal of $3.1 = .3225$. The confidence interval is then given by

$$\begin{aligned} 1-\alpha &= \Pr\left[\frac{565\,000}{362\,500}\frac{1}{3.1} \leq \frac{\sigma_c^2}{\sigma_p^2} \leq \frac{565\,000}{362\,500}\frac{1}{.297}\right] \\ &= \Pr\left[.502\,78 \leq \frac{\sigma_c^2}{\sigma_p^2} \leq 5.248\right] \end{aligned} \tag{34}$$

For problems such as this, it is much easier to use computer programs to obtain the necessary values and probabilities from the F distribution than to use tables.

7. An example using data on exam grades

Consider the following scores from a graduate economics class which has eighteen students.

Scores $= \{46,\ 58,\ 87,\ 97.5,\ 82.5,\ 68,\ 83.25,\ 99.5,\ 66.5,\ 75.5,\ 62.5,\ 67,\ 78,\ 32,\ 74.5,\ 47,\ 99.5,\ 26\}$

The mean of the data is 69.4583. The variance is 466.899 and the standard deviation is 21.6078. We are interested in the null hypothesis $\mu = 80$. If we look at the tabled distribution of the t distribution with 17 degrees of freedom and with 0.025 in each tail we see that $\gamma_1 = 2.110$. This means then that

$$
\begin{aligned}
1-\alpha = 0.95 = F(\gamma_1) - F(-\gamma_1) &= \Pr\left[-\gamma_1 \leq \frac{\sqrt{n}(\bar{y}-\mu)}{S} \leq \gamma_1\right] \\
&= \Pr\left[-2.110 \leq \frac{4.2426(69.4583-\mu)}{21.6078} \leq 2.110\right] \\
&= \Pr\left[-10.746\,35 \leq 69.4583 - \mu \leq 10.746\,35\right] \\
&= \Pr\left[10.746\,35 \geq -69.4583 + \mu \geq -10.746\,35\right] \\
&= \Pr\left[80.204\,65 \geq 69.4583 + \mu \geq -58.7119\right] \\
&= \Pr\left[58.7119 \leq \mu \leq 80.204\,65\right]
\end{aligned} \tag{35}
$$

Given that 80 lies within this bound, we cannot reject the hypothesis that $\mu = 80$.

8. Hypothesis Testing

8.1. A statistical hypothesis is an assertion or conjecture about the distribution of one or more random variables. If a statistical hypothesis completely specifies the distribution, it is referred to as a **simple hypothesis**; if not, it is referred to as a **composite hypothesis**. A simple hypothesis must therefore specify not only the functional form of the underlying distribution, but also the values of all parameters. A statement regarding a parameter θ, such as $\theta \in \omega \subset \Omega$, is called a statistical hypothesis about θ and is usually referred to by H_0. The decision on accepting or rejecting the hypothesis is based on the value of a certain random variable or test statistic X, the distribution P_θ of which is known to belong to a class $\mathcal{P} = \{P_\theta, , \theta \in \Omega\}$. We shall assume that is θ were known, one would also known whether the hypothesis were true. The distributions of $\mathcal{P}$ can then be classified into those for which the hypothesis is ture and those for which it is false. The resulting two mutually exclusive classes are denoted H and K (or H_0 and H_1), and the corresponding subsets of Ω by Ω_H and Ω_K respectively, so that $H \cup K = \mathcal{P}$ and $\Omega_H \cup \Omega_K = \Omega$. Mathematically the hypothesis is equivalent to the statement that P_θ is an element of H.

8.2. Alternative hypotheses. To test statistical hypotheses it is necessary to formulate alternative hypotheses. The statement that $\theta \in \omega^c$ (the complement of ω with respect to Ω) is also a statistical hypothesis about θ, which is called the *alternative* to H_0 and is usually denoted H_1 or H_A. Thus we have

$$H_0 : \theta \in \omega$$

$$H_1 : \theta \in \omega^c$$

If ω contains only one point, that is, $\omega = \{\theta_0\}$, then H_0 is called a *simple* hypothesis, otherwise it is called a *composite* hypothesis. A composite hypothesis H_0 might be that mean of a population μ is greater than or equal to _0, i.e.,

$$\mu \geq \mu_0$$

with composite alternative

$$\mu < \mu_0$$

We usually call the hypothesis we are testing the **null** hypothesis. The name comes from experiments where θ measures the response to some treatment and H_0:$\theta_0 = 0$ whereas the alternative is H_1:$\theta_0 \neq 0$. The null hypothesis is then that the treatment has no effect. Frequently statisticians state their hypotheses as the opposite of what they believe to be true with the hope that the test procedures will lead to their rejection.

8.3. Acceptance and rejection regions. Let the decision of accepting or rejecting H_0 be denoted by d_0 and d_1 respectively. A nonrandomized test procedure assigns to each possible x of X one of these two decisions and the thereby divides the sample space into two complementary regions S_0 and S_1. If X falls into S_0 the hypothesis is not rejected; otherwise it is rejected. The set S_0 is called the region of acceptance, the set S_1 is called the region of rejection or **critical** region.

The typical procedure to test hypotheses is to compute a test statistic X which will tell whether the hypothesis should be accepted or rejected. The sample space is partitioned into two regions: the acceptance region and the rejection region. If X is in the rejection region the hypothesis is rejected, and not rejected otherwise.

8.4. Tests where H_0 is of the form $\theta = \theta_0$ and the alternative H_1 is of the two-sided form $\theta \neq \theta_0$ are called two-tailed tests. When H_0 is of the form $\theta = \theta_0$ but the alternative is H_1: $\theta < \theta_0$ or H_1: $\theta > \theta_0$, the test is called a one-tailed test.

8.5. Type I and Type II errors. Suppose that the null hypothesis is $\theta = \theta_0$ and the alternative is $\theta = \theta_1$. The statistician can make two possible types of errors. If $\theta = \theta_0$ and the test rejects $\theta = \theta_0$ and concludes $\theta = \theta_1$ then the error is of type I.

Rejection of the null hypothesis when it is true is a type I error.

If $\theta = \theta_1$ and the test does not reject $\theta = \theta_0$ but accepts that $\theta = \theta_0$ then the error is of type II.

Acceptance of the null hypothesis when it is false is a type II error.

Suppose that the null hypothesis is that a person tested for HIV does not have the disease. Then a false positive is a Type I error while a false negative is a Type II error. We can summarize as follows where S_1 is the region of rejection or **critical** region.

$$P_\theta\{X \in S_1\} = \text{Probability of a Type I error if } \theta \in \Omega_H \tag{36a}$$

$$P_\theta\{X \in S_1\} = (\text{1 - Probability of a Type II error}) \text{ if } \theta \in \Omega_K \tag{36b}$$

$$P_\theta\{X \in S_0\} = \text{Probability of a Type II error if } \theta \in \Omega_K \tag{36c}$$

8.6. Level of significance, size and power. It is customary to assign a bound to the probability of incorrectly rejecting H_0 when it is true (a Type I error) and attempt to minimize the probability of accepting H_0 when it is false (a Type II error) subject to the condition on the Type I error. Thus one selects a number α between 0 and 1, called the level of significance, and imposes the condition that

$$P_\theta\{\delta(X) = d_1\} = P_\theta\{X \in S_1\} \leq \alpha, \text{ for all } \theta \in \Omega_H \tag{37}$$

where $\delta(\cdot)$ is an indicator choosing d_0 or d_1. This says that the probability that $X \in S_1$ is less than α when θ is in the hypothesized part of its domain.

Subject to the condition in equation 37, it is desired to either

$$\min_{\theta \in \Omega_K} P_\theta\{\delta(X) = d_0\} = \min_{\theta \in \Omega_K} P_\theta\{X \in S_0\} \tag{38}$$

or

$$\max_{\theta \in \Omega_K} P_\theta\{\delta(X) = d_1\} = \max_{\theta \in \Omega_K} P_\theta\{X \in S_1\} \tag{39}$$

Although 39 usually implies that

$$\sup_{\theta \in \Omega_H} P_\theta\{X \in S_1\} = \alpha \tag{40}$$

it is convenient to introduce a term for the left hand side of equation 40; it is called the size of the test or critical region S_1.

$$\text{Size of the test} = \sup_{\theta \in \Omega_H} P_\theta\{X \in S_1\} \tag{41}$$

We then say for $0 \leq \alpha \leq 1$

$$\begin{aligned} &\text{If } \sup_{\theta \in \Omega_H} P_\theta\{X \in S_1\} \; = \; \alpha, \text{ the test is a } \textbf{size } \alpha \text{ test} \\ &\text{If } \sup_{\theta \in \Omega_H} P_\theta\{X \in S_1\} \; \leq \; \alpha, \text{ the test is a } \textbf{level } \alpha \text{ test} \end{aligned} \tag{42}$$

Equation 37 restricts consideration to tests whose size does not exceed α or which have a level equal to α. The probability of rejection (equation 39) evaluated for a given $\theta \in \Omega_K$ is called the **power** of the test against the alternative θ. Considered as a function of θ for all $\theta \in \Omega$, the probability in equation 39 is called the power function of the test and is denoted by $\beta(\theta)$.

$$\text{Power Function} \; = \; \beta(\theta) \; = \; P_\theta\{X \in S_1\} \tag{43}$$

The probability of a Type II error is given by the probability in equation 38 or by 1 minus the probability in equation 39. Note that in either case we case considering either $\beta(\theta)$ or 1 - $\beta(\theta)$ where θ is now restricted to $\{\theta \in \Omega_K\}$.

In some ways, the size of the critical region is just the probability of committing a Type I error. And we loosely call this probability of committing a Type I error the level of significance of the test. The probability of committing a type II error is denoted by 1-$\beta(\theta|\theta \in \Omega_K)$. The ideal power function is 0 for all $\theta \in \Omega_H$ and 1 for all $\theta \in \Omega_K$. Except in trivial situations, this ideal cannot be attained. Qualitatively, a good test has power function near 1 for most $\theta \in \Omega_K$ and near 0 for most $\theta \in \Omega_H$.

8.7. The probabilities of committing the two types of error can be summarized as follows.

		Decision	
		Do not reject H_0	Reject H_0
$\boldsymbol{H_0}$	True	Correct Decision	Type I Error
	False	Type II Error	Correct Decision

Because we do not know in fact whether H_0 or H_1 is ture, we cannot tell whether a test has made a correct decision or not. If a test rejects H_0, we do not know whether the decision is correct or a Type I error is made. If a test fails to reject H$-$0, we do not know whether a correct decision is made or a Type II error is made. We can only assess the long term accuracy of decisions made by a test by adopting a frequentist approach as we did with confidence intervals.

Consider a long series of trials in which the same decision rule is applied to repeated samples drawn under identical conditions (either H_0 or under H_1). The limiting proportion of trials in which H_0 is rejected when H_0 is true is the **probability of a type I error**, the α**-risk**. Similarly, the limiting proportion of trials in which H_0 is not rejected when H_1 is true is the **probability of a type II error**, also called the 1-β**-risk**. This frequentist interpretation of error probabilities leads to the following definitions:

$$\begin{aligned} \alpha \; = \; P\{\text{Type I error}\} \; &= P\{\text{Reject H}_0 \text{ when H}_0 \text{ is true}\} \\ &= P\{\text{Reject H}_0|\text{H}_0\} \end{aligned} \tag{44}$$

and

$$\begin{aligned} 1 - \beta \; = \; P\{\text{Type II error}\} \; &= P\{\text{Fail to reject H}_0 \text{ when H}_1 \text{ is true}\} \\ &= P\{\text{Fail to reject H}_0|\text{H}_1\} \end{aligned} \tag{45}$$

The power of the test, of course, is

$$\beta(\theta|\theta \in \Omega_K) = P\{\text{Reject H}_0 \mid H_1\} \tag{46}$$

The **power** of the test, is the probability that the test correctly rejects H_0. Keeping in mind that H_1 is the research hypothesis to be proved, the power of a test measures its ability to "prove" H_1 when H_1 is true. A test with high power is preferred.

8.8. Probability values (*p*-values) for statistical tests. The p-value associated with a statistical test is the probability that we obtain the observed value of the test statistic or a value that is more extreme in the direction of the alternative hypothesis calculated when H_0 is true. Rather than specifying the critical region ahead of time, the p-value of a test can be reported and the reader make a decision based on it.

If T is a test statistic, the p-value, or *attained significance level*, is the smallest level of significance α for which the observed data indicate that the null hypothesis should be rejected.

The smaller the p-value becomes, the more compelling is the evidence that the null hypothesis should be rejected. If the desired level of significance for a statistical test is greater than or equal to the p-value, the null hypothesis is rejected for that value of α. The null hypothesis should be rejected for any value of α down to and including the p-value.

To compute a p-value, we compute the realized value of the test statistic for a given sample and then find the probability that the random variable represented by the test statistic is larger than this realized value in the given sample. If a test statistic happened to be distributed N(0,1), we would simply find the probability that a standardized normal random variable was larger than realized value of the test statistic. For example, if the realized value of the test statistic were 1.96 for a one sided test, we would conclude that the probability we would reject the null hypothesis was 0.975 or a p-value of 0.025.

8.9. Likelihood ratio tests.

8.9.1. *The general approach to likelihood ratio tests.* Let X_1, X_2, ..., Xn be a random sample from a population with pdf f(x | θ). Then the likelihood function of a sample is given by

$$L(\theta|x_1, x_2, \ldots, x_n) = L(\theta|x) = \prod_{i=1}^{n} f(x_i|\theta) \tag{47}$$

Let Ω denote the entire parameter space such that $\theta \in \Omega$. The *likelihood ratio test statistic* for testing $\{H_0 : \theta \in \Omega_0\}$ versus $\{H_1 : \theta \in \Omega_1\}$ is

$$\lambda(x) = \frac{\sup\limits_{\Omega_0} L(\theta|x)}{\sup\limits_{\Omega} L(\theta|x)} \tag{48}$$

A *likelihood ratio test* (LRT) is any test that has a rejection region of the form $\{x : \lambda(x) \leq c\}$ where c is any number satisfying $0 \leq c \leq 1$. The numerator in 48 is the maximum of the likelihood if θ is restricted to Ω_0 while the denominator is the maximum of the likelihood if θ is allowed to vary over the entire parameter space. $\lambda(x)$ is small if there are parameter points in the alternative hypothesis Ω_1 for which the observed sample is much more likely than for any parameter point in the null hypothesis.

If we think of doing the maximization over both the entire parameter space (unrestricted maximization) and a subset of the parameter space (restricted maximization), then a clear correspondence between maximum likelihood estimation (MLE) and likelihood ratio tests (LRT) becomes apparent. Suppose $\hat{\theta}$, an MLE of θ, exists and $\hat{\theta}$ is obtained by maximizing $L(\theta|x)$ without any restrictions on Ω. Then consider the MLE of θ_0, call it $\hat{\theta}_0$ which is obtained by maximizing $L(\theta|x)$ subject to the constraint that $\theta \in \Omega_0$. That is, $\hat{\theta}_0 = \hat{\theta}_0(x)$ is the value of $\theta \in \Omega_0$ that maximizes $L(\theta|x)$. Then the LRT statistic is given by

$$\lambda(x) \;=\; \frac{L(\hat{\theta}_0|x)}{L(\hat{\theta}|x)} \tag{49}$$

We reject the null hypothesis that $H_0 : \theta = \theta_0$ if

$$\lambda(x) \;=\; \frac{L(\hat{\theta}_0|x)}{L(\hat{\theta}|x)} \;<\; k \tag{50}$$

where k is some constant. In general we do not know the distribution of $\lambda(x)$. Common practice is to work with $-\log \lambda(x)$ for which the critical region becomes

$$\begin{aligned} \log\left[\lambda(x)\right] \;&=\; \log\left[L(\hat{\theta}_0|x)\right] \;-\; \log\left[L(\hat{\theta}|x)\right] \;<\; \log k \\ \Rightarrow\; -\log\left[\lambda(x)\right] \;&=\; \log\left[L(\hat{\theta}|x)\right] \;-\; \log\left[L(\hat{\theta}_0|x)\right] \;>\; -\log k \\ \Rightarrow\; -\log\left[\lambda(x)\right] \;&>c, \quad c \;=\; -\log k \;>\; 0 \text{ because k} \leq 1 \end{aligned} \tag{51}$$

Asymptotically, $-2\log\left[\lambda(x)\right]$ is distributed as a χ^2 random variable with degrees of freedom equal to the difference in the dimension of Ω and Ω_H, which is often equal to the number of restrictions imposed by the null hypothesis.

8.9.2. *Example likelihood ratio test for the mean with a known variance.* Consider a random sample X_1, X_2, . . . , Xn from a N(μ,σ^2) population. Consider testing $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$. We know from previous examples that the MLE estimator of μ is $\bar{X}$. Thus the denominator of $\lambda(x)$ is L($\bar{x}|x_1, x_2, \ldots, x_n$). So the LRT statistic is

$$\begin{aligned} \lambda(x) \;&=\; \frac{\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \; e^{\frac{-1}{2\sigma^2} \; \sum_{i=1}^n (x_i - \mu_0)^2}}{\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \; e^{\frac{-1}{2\sigma^2} \; \sum_{i=1}^n (x_i - \bar{x})^2}} \\ &= \frac{e^{\frac{-1}{2\sigma^2} \; \sum_{i=1}^n (x_i - \mu_0)^2}}{e^{\frac{-1}{2\sigma^2} \; \sum_{i=1}^n (x_i - \bar{x})^2}} \\ &= e^{\frac{-1}{2\sigma^2} \left(\sum_{i=1}^n (x_i - \mu_0)^2 \;-\; \sum_{i=1}^n (x_i - \bar{x})^2\right)} \end{aligned} \tag{52}$$

Notice that we can write $\sum_{i=1}^n (x_i \;-\; \mu_0)^2$ as follows

$$\begin{aligned}\sum_{i=1}^{n}(x_i - \mu_0)^2 &= \sum_{i=1}^{n}(x_i - \bar{x} + \bar{x} - \mu_0)^2 \\ &= \sum_{i=1}^{n}(x_i - \bar{x})^2 + 2(\bar{x} - \mu_0)\sum_{i=1}^{n}(x_i - \bar{x}) + n(\bar{x} - \mu_0)^2 \\ &= \sum_{i=1}^{n}(x_i - \bar{x})^2 + 2(\bar{x} - \mu_0)\,(n\bar{x} - n\bar{x}) + n(\bar{x} - \mu_0)^2 \\ &= \sum_{i=1}^{n}(x_i - \bar{x})^2 + n(\bar{x} - \mu_0)^2\end{aligned} \tag{53}$$

Now substitute for $\sum_{i=1}^{n}(x_i - \mu_0)^2$ in equation 52 from equation 53

$$\begin{aligned}\lambda(x) &= e^{\frac{-1}{2\sigma^2}\left(\sum_{i=1}^{n}(x_i - \mu_0)^2 - \sum_{i=1}^{n}(x_i - \bar{x})^2\right)} \\ &= e^{\frac{-1}{2\sigma^2}\left(\sum_{i=1}^{n}(x_i - \bar{x})^2 + n(\bar{x} - \mu_0)^2 - \sum_{i=1}^{n}(x_i - \bar{x})^2\right)} \\ &= e^{\frac{-n}{2\sigma^2}\,(\bar{x} - \mu_0)^2}\end{aligned} \tag{54}$$

An LRT test will reject $H_0 : \mu = \mu_0$ for **"small"** values of $\lambda(x)$. That is, reject $H_0 : \mu = \mu_0$ if

$$\lambda(x) = e^{\frac{-n}{2\sigma^2}\,(\bar{x} - \mu_0)^2} \le k \tag{55a}$$

$$\Rightarrow \frac{-n}{2\sigma^2}\,(\bar{x} - \mu_0)^2 \le \log k \tag{55b}$$

$$\Rightarrow (\bar{x} - \mu_0)^2 \ge \frac{-2\sigma^2}{n}\log k \tag{55c}$$

$$\Rightarrow |\bar{x} - \mu_0| \ge \sqrt{\frac{-2\sigma^2}{n}\log k} \tag{55d}$$

$$\Rightarrow \frac{|\bar{x} - \mu_0|}{\frac{\sigma}{n}} \ge \frac{\sqrt{\frac{-2\sigma^2}{n}\log k}}{\frac{\sigma}{n}} \tag{55e}$$

$$\Rightarrow \frac{|\bar{x} - \mu_0|}{\frac{\sigma}{n}} \ge \gamma \tag{55f}$$

where γ will be determined so that the critical region is of the appropriate size. Note that $\log k$ is negative in view of the fact that 0 < k <1. Because $\bar{X} \sim N\left(\mu_0, \frac{\sigma^2}{n}\right)$ under the null hypothesis, we can use the standard normal table to create an appropriate critical level. Specifically, we reject $H_0 : \mu = \mu_0$ if

$$\left|\frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}\right| \ge Z_{\alpha/2} \tag{56}$$

Now consider $-\log\lambda(x)$ in equation 55

$$\lambda(x) = e^{\frac{-n}{2\sigma^2}\,(\bar{x} - \mu_0)^2} \leq k \tag{57a}$$

$$\Rightarrow\ 2\log\lambda(x) = 2\left(\frac{-n}{2\sigma^2}\right)(\bar{x} - \mu_0)^2 \leq \log k \tag{57b}$$

$$\Rightarrow\ -\log\lambda(x) = \frac{n}{\sigma^2}(\bar{x} - \mu_0)^2 \geq -\log k \tag{57c}$$

$$= \frac{(\bar{x} - \mu_0)^2}{\frac{\sigma^2}{n}} \geq c, \quad c = -\log k > 0 \text{ because } k \leq 1 \tag{57d}$$

$$= \left(\frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}\right)^2 \geq c \tag{57e}$$

The expression on the left hand side of equation 57e is of course the square of a standard normal varible which is distributed as a χ^2 random variable with one degree of freedom. So for this case the asymptotic result holds exactly in all samples.

8.9.3. *Example likelihood ratio test for the mean and unknown variance.* Consider a random sample X_1, $X_2, \ldots, Xn$ from a N(μ,σ^2) population. Consider testing $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$. We know from previous examples that the MLE estimator of μ is $\bar{X}$ and the unrestricted MLE estimator of σ^2 is $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$. The restricted estimators are $\hat{\mu}_0 = \mu_0$ and $\hat{\sigma}_0^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu_0)^2$. Summarizing

$$\begin{aligned}
\hat{\mu}_0 &= \mu_0 \\
\hat{\sigma}_0^2 &= \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu_0)^2 \\
\hat{\mu} &= \bar{x} \\
\hat{\sigma}^2 &= \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2
\end{aligned} \tag{58}$$

So the LRT statistic is

$$\lambda(x) = \frac{\left(\frac{1}{\sqrt{2\pi}\hat{\sigma}_0}\right)^n e^{\frac{-1}{2\hat{\sigma}_0^2}\sum_{i=1}^{n}(x_i - \mu_0)^2}}{\left(\frac{1}{\sqrt{2\pi}\hat{\sigma}}\right)^n e^{\frac{-1}{2\hat{\sigma}^2}\sum_{i=1}^{n}(x_i - \bar{x})^2}} \tag{59}$$

Now rewrite the numerator in equation 59 as follows

$$\begin{aligned}
\left(\frac{1}{\sqrt{2\pi}\hat{\sigma}_0}\right)^n e^{\frac{-1}{2\hat{\sigma}_0^2}\sum_{i=1}^{n}(x_i - \mu_0)^2} &= \left(\frac{1}{\sqrt{2\pi}\hat{\sigma}_0}\right)^n e^{\frac{-n\hat{\sigma}_0^2}{2\hat{\sigma}_0^2}} \\
&= \left(\frac{1}{\sqrt{2\pi}\hat{\sigma}_0}\right)^n e^{\frac{-n}{2}}
\end{aligned} \tag{60}$$

Then rewrite the denominator in equation 59 as follows

$$
\begin{aligned}
\left(\frac{1}{\sqrt{2\pi}\hat{\sigma}}\right)^n e^{\frac{-1}{2\hat{\sigma}^2}\sum_{i=1}^n (x_i - \bar{x})^2} &= \left(\frac{1}{\sqrt{2\pi}\hat{\sigma}}\right)^n e^{\frac{-n\hat{\sigma}^2}{2\hat{\sigma}^2}} \\
&= \left(\frac{1}{\sqrt{2\pi}\hat{\sigma}}\right)^n e^{\frac{-n}{2}}
\end{aligned}
\tag{61}
$$

Now write equation 59 as follows

$$
\begin{aligned}
\lambda(x) &= \frac{\left(\frac{1}{\sqrt{2\pi}\hat{\sigma}_0}\right)^n e^{\frac{-1}{2\hat{\sigma}_0^2}\sum_{i=1}^n (x_i - \mu_0)^2}}{\left(\frac{1}{\sqrt{2\pi}\hat{\sigma}}\right)^n e^{\frac{-1}{2\hat{\sigma}^2}\sum_{i=1}^n (x_i - \bar{x})^2}} \\
&= \frac{\left(\frac{1}{\sqrt{2\pi}\hat{\sigma}_0}\right)^n}{\left(\frac{1}{\sqrt{2\pi}\hat{\sigma}}\right)^n} = \left(\frac{\frac{1}{\hat{\sigma}_0}}{\frac{1}{\hat{\sigma}}}\right)^n = \left(\frac{\hat{\sigma}_0}{\hat{\sigma}}\right)^{-n}
\end{aligned}
\tag{62}
$$

Now substitute from equation 58 into equation 62 as follows

$$
\begin{aligned}
\lambda(x) &= \left(\frac{\hat{\sigma}_0}{\hat{\sigma}}\right)^{-n} \\
&= \left(\frac{\sum_{i=1}^n (x_i - \mu_0)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}\right)^{\frac{-n}{2}}
\end{aligned}
\tag{63}
$$

Now make the substitution for $\sum_{i=1}^n (x_i - \mu_0)^2$ from equation 53 into equation 63

$$
\begin{aligned}
\lambda(x) &= \left(\frac{\sum_{i=1}^n (x_i - \mu_0)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}\right)^{\frac{-n}{2}} \\
&= \left(\frac{\sum_{i=1}^n (x_i - \bar{x})^2 + n(\bar{x} - \mu_0)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}\right)^{\frac{-n}{2}} \\
&= \left(1 + \frac{n(\bar{x} - \mu_0)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}\right)^{\frac{-n}{2}}
\end{aligned}
\tag{64}
$$

An LRT test will reject $H_0 : \mu = \mu_0$ for **"small"** values of $\lambda(x)$. That is, reject $H_0 : \mu = \mu_0$ if

$$\lambda(x) = \left(1 + \frac{n(\bar{x} - \mu_0)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)^{\frac{-n}{2}} \leq c \tag{65a}$$

$$\Rightarrow \left(1 + \frac{n(\bar{x} - \mu_0)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right) \geq c^{\frac{-2}{n}} \tag{65b}$$

$$\Rightarrow \frac{n(\bar{x} - \mu_0)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \geq c^{\frac{-2}{n}} - 1 \tag{65c}$$

$$\Rightarrow \left(\frac{\sqrt{n}(\bar{x} - \mu_0)}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}}\right)^2 \geq c^{\frac{-2}{n}} - 1 \tag{65d}$$

$$\Rightarrow \frac{1}{n-1}\left(\frac{\sqrt{n}(\bar{x} - \mu_0)}{\sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}}\right)^2 \geq c^{\frac{-2}{n}} - 1 \tag{65e}$$

$$\Rightarrow \frac{1}{n-1}\left(\frac{\sqrt{n}(\bar{x} - \mu_0)}{s}\right)^2 \geq c^{\frac{-2}{n}} - 1 \tag{65f}$$

$$\Rightarrow \left(\frac{\sqrt{n}(\bar{x} - \mu_0)}{s}\right)^2 \geq (n-1)\left(c^{\frac{-2}{n}} - 1\right) \tag{65g}$$

$$\Rightarrow \left(\frac{\sqrt{n}(\bar{x} - \mu_0)}{s}\right) \geq \sqrt{(n-1)\left(c^{\frac{-2}{n}} - 1\right)} = \gamma \tag{65h}$$

where γ will be determined so that the critical region is of the appropriate size. Note that the left hand side of equation 65h is distributed as a t random variable from equation 17 where we substitute $\bar{x}$ for $\hat{\beta}$ and μ_0 for β. Specifically, we reject $H_0 : \mu = \mu_0$ if

$$\left|\frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}\right| \geq t_{\alpha/2}(n-1) \tag{66}$$

8.9.4. *More on the classroom grade example.* Consider the following scores from a graduate economics class which has eighteen students.

Scores $= \{46,\ 58,\ 87,\ 97.5,\ 82.5,\ 68,\ 83.25,\ 99.5,\ 66.5,\ 75.5,\ 62.5,\ 67,\ 78,\ 32,\ 74.5,\ 47,\ 99.5,\ 26\}$

The mean of the data is 69.4583. The variance is 466.899 and the standard deviation is 21.6078. We are interested in the null hypothesis $\mu = 80$. We can compute the t-statistic as follows.

$$\begin{aligned} t &= \frac{\bar{x} - \mu}{\frac{S}{\sqrt{n}}} \\ &= \frac{69.4583 - 80}{\frac{21.6078}{\sqrt{18}}} \\ &= -2.069\,83 \end{aligned} \tag{67}$$

Consider the null hypothesis that $\mu = 80$. Consider an α level of $\alpha = 0.05$. The value of the t statistic is $-2.069\,83$. We reject the null hypothesis if $|-2.069\,83| > t_{n-1,\,.025}$. From the t-tables $t_{17,\,.025} = 2.110$. Given that 2.069 83 is less than 2.110, we cannot reject the null hypothesis that $\mu = 80$.

8.10. Most Powerful Tests.

A test with level α that has the highest possible power is called a **most powerful** test. We abbreviate most powerful with MP and say that a test statistic is MP if it is associated with an MP test.

8.10.1. *Test functions.*

Consider a test statistic X. For the likelihood ratio test, X = $\lambda(x)$. In equation 37 we used an indicator function to determine whether a given value of X implied we should accept or reject the null hypothesis. Specifically, we choose $\delta(X)$ such that.

$$\begin{aligned} &\text{If } \delta(X) \ = d_0, \text{ then } X \in S_0 \text{ and we accept H}_0 \\ &\text{If } \delta(X) \ = d_1, \text{ then } X \in S_1 \text{ and we reject H}_0 \end{aligned} \tag{68}$$

Now consider a different indicator function $\varphi_k(x)$ where $0 \le k \le \infty$ and x is the observation vector. We define $\varphi_k(x)$ as follows

$$\varphi(x) = \begin{cases} 1 & \text{if } X > k \\ 0 & \text{if } X < k \end{cases} \tag{69}$$

with $\varphi_k(X)$ any value in (0,1) if equality occurs. The value of k is such that $X > k \Rightarrow X \in S_1$. For the likelihood ratio test this would yield

$$\varphi(x) = \begin{cases} 1 & \text{if } \lambda(x, \theta_0, \theta_1) > k \\ 0 & \text{if } \lambda(x, \theta_0, \theta_1) < k \end{cases} \tag{70}$$

Because we want the results on power valid for all possible test sizes $\alpha \in (0,1)$, we consider *randomized tests* φ, which are tests that may take values in (0,1) as compared to taking values of 0 or 1 unless X = k. If $0 < \varphi(x) < 1$ for the observation vector x, the interpretation is that we toss a coin with probability of heads $\varphi(x)$ and reject H_0 iff the coin shows heads. Such randomized tests are not used in practice. They are only used to show that with randomization, the likelihood ratio tests are most powerful no matter what size we choose for α.

8.10.2. *Neyman-Pearson Lemma.*

Theorem 1 (Neyman-Pearson Lemma)**.**

(a) *If $\alpha > 0$ and φ_k is a size α likelihood ratio test, then φ_k is MP in the class of level α tests.*
(b) *For each $0 \le \alpha \le 1$ there exists an MP size α likelihood ratio test provided that randomization is permitted, $0 \le \varphi \le 1$, for some x.*
(c) *If φ is an MP level α test, then it must be a level α likelihood ratio test; that is there exists k such that*

$$P_\theta\left[\varphi(\cdot) \neq \varphi_k(\cdot), \lambda(\cdot, \theta_0, \theta_1) \neq k\right] = 0 \tag{71}$$

for $\theta = \theta_0$ and $\theta = \theta_1$.

9. Summary Tables on Common Statistical Tests

Tables 5, 6, 7, 8, 9, 10 and 11 below summarize information on conducting some of the more common hypothesis tests on simple random samples. We use the notation that

$$z_\alpha = \left[z_\alpha : \int_{z_\alpha}^{\infty} f(z;\, \theta)\, dz = \alpha \right]$$

i.e. z_α is the value such that α of the distribution lies to the right of z_α.

Table 5. Level α Tests on μ when σ^2 is Known.
Test Statistic: $\mathbf{z} = \dfrac{\bar{\mathbf{x}} - \mu_0}{\sigma/\sqrt{\mathbf{n}}}$.
($\Phi(z)$ is the distribution function for a standard normal random variable.)

Testing Problem	Hypotheses	Reject H_0 if	P-Value
Upper One-sided	$H_0: \mu \leq \mu_0$ vs $H_1: \mu > \mu_0$	$z > z_\alpha$ $\Leftrightarrow$ $\bar{x} > \mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}}$	$P(Z \geq z \mid H_0)$ = $1 - \Phi(z)$
Lower One-sided	$H_0: \mu \geq \mu_0$ vs $H_1: \mu < \mu_0$	$z < -z_a$ $\Leftrightarrow$ $\bar{x} < \mu_0 - z_\alpha \frac{\sigma}{\sqrt{n}}$	$P(Z \leq z \mid H_0)$ = $\Phi(z)$
Two-sided	$H_0: \mu = \mu_0$ vs $H_1: \mu \neq \mu_0$	$\lvert z \rvert > z_{\alpha/2}$ $\Leftrightarrow$ $\lvert \bar{x} - \mu_0 \rvert > z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$	$P(\lvert Z \rvert \geq \lvert z \rvert \mid H_0)$ = $2[1 - \Phi(\lvert z \rvert)]$

TABLE 6. Level α Tests on μ when σ^2 is Unknown.
Test Statistic: $\mathbf{t} = \dfrac{\bar{\mathbf{x}} - \mu_0}{\mathbf{S}/\sqrt{\mathbf{n}}}$.
($P(T_{n-1} \geq t)$ is the area in the upper tail of the t-distribution greater than t.)

Testing Problem	Hypotheses	Reject H_0 if	P-Value
Upper One-sided	$H_0: \mu \leq \mu_0$ vs $H_1: \mu > \mu_0$	$t > t_{n-l,\,\alpha}$ $\Leftrightarrow$ $\bar{x} > \mu_0 + t_{n-1,\,\alpha}\dfrac{S}{\sqrt{n}}$	$P(T_{n-1} \geq t)$
Lower One-sided	$H_0: \mu \geq \mu_0$ vs $H_1: \mu < \mu_0$	$t < -t_{n-1,\,\alpha}$ $\Leftrightarrow$ $\bar{x} < \mu_0 - t_{n-1,\,\alpha}\dfrac{S}{\sqrt{n}}$	$P(T_{n-1} \leq t)$
Two-sided	$H_0: \mu = \mu_0$ vs $H_1: \mu \neq \mu_0$	$\lvert t\rvert > t_{n-1,\,\alpha/2}$ $\Leftrightarrow$ $\lvert\bar{x} - \mu_0\rvert > t_{n-1,\,\alpha/2}\dfrac{S}{\sqrt{n}}$	$P(\lvert T_{n-1}\rvert \geq \lvert t\rvert)$ = $2P(T_{n-1} \geq \lvert t\rvert)$

TABLE 7. Level α Tests on σ^2.
Test Statistic: $\chi^{\mathbf{2}} = \dfrac{(\mathbf{n-1})\mathbf{s^2}}{\sigma_{\mathbf{0}}^{\mathbf{2}}}$
($P_U = P\left(\chi^2_{n-1} \geq \chi^2\right)$ is the area in the upper tail of the χ^2_{n-1} distribution greater than χ^2.)

Testing Problem	Hypotheses	Reject H_0 if	P-Value
Upper One-sided	$H_0: \sigma^2 \leq \sigma_0^2$ vs $H_1: \sigma^2 > \sigma_0^2$	$\chi^2 > \chi^2_{n-1,\,\alpha}$ $\Leftrightarrow$ $s^2 > \dfrac{\sigma_0^2\chi^2_{n-1,\,\alpha}}{n-1}$	$P_U = P\left(\chi^2_{n-1} \geq \chi^2\right)$
Lower One-sided	$H_0: \sigma^2 \geq \sigma_0^2$ vs $H_1: \sigma^2 < \sigma_0^2$	$\chi^2 < \chi^2_{n-1,\,1-\alpha}$ $\Leftrightarrow$ $s^2 < \dfrac{\sigma_0^2\chi^2_{n-1,\,1-\alpha}}{n-1}$	$P_L = P\left(\chi^2_{n-1} \leq \chi^2\right)$
Two-sided	$H_0: \sigma^2 = \sigma_0^2$ vs $H_1: \sigma^2 \neq \sigma_0^2$	$\chi^2 > \chi^2_{n-1,\,\alpha/2}$ or $\chi^2 < \chi^2_{n-1,\,1-\alpha/2}$ $\Leftrightarrow$ $s^2 > \dfrac{\sigma_0^2\chi^2_{n-1,\,\alpha/2}}{n-1}$ or $s^2 < \dfrac{\sigma_0^2\chi^2_{n-1,\,1-\alpha/2}}{n-1}$	$2\min\{P_U,\, P_L = 1 - P_U\}$

10. MORE ON THE CLASSROOM GRADE EXAMPLE

Consider the following scores from a graduate economics class which has eighteen students.

$$\text{Scores} = \{46,\ 58,\ 87,\ 97.5,\ 82.5,\ 68,\ 83.25,\ 99.5,\ 66.5,\ 75.5,\ 62.5,\ 67,\ 78,\ 32,\ 74.5,\ 47,\ 99.5,\ 26\}$$

TABLE 8. Level α Tests on $\mu_1 - \mu_2$ for Large Samples.
Test Statistic: $\mathbf{z} = \dfrac{\bar{\mathbf{x}} - \bar{\mathbf{y}} - \delta_0}{\sqrt{\frac{\mathbf{s_1^2}}{\mathbf{n_1}} + \frac{\mathbf{s_2^2}}{\mathbf{n_2}}}}$.
($P(Z \geq z)$ is the area in the upper tail of the standard normal distribution greater than z.)

Testing Problem	Hypotheses	Reject H_0 if	P-Value
Upper One-sided	$H_0: \mu_1 - \mu_2 \leq \delta_0$ vs $H_1: \mu_1 - \mu_2 > \delta_0$	$z > z_\alpha$ $\Leftrightarrow$ $\bar{x} - \bar{y} > \delta_0 + z_\alpha\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$	$P(Z \geq z)$
Lower One-sided	$H_0: \mu_1 - \mu_2 \geq \delta_0$ vs $H_1: \mu_1 - \mu_2 < \delta_0$	$z < -z_\alpha$ $\Leftrightarrow$ $\bar{x} - \bar{y} < \delta_0 - z_\alpha\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$	$P(Z \leq z)$
Two-sided	$H_0: \mu_1 - \mu_2 = \delta_0$ vs $H_1: \mu_1 - \mu_2 \neq \delta_0$	$\lvert z\rvert > z_{\alpha/2}$ $\Leftrightarrow$ $\lvert\bar{x} - \bar{y} - \delta_0\rvert > z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$	$P(\lvert Z\rvert \geq \lvert z\rvert)$ = $2[1 - \Phi(\lvert z\rvert)]$

TABLE 9. Level α Tests on $\mu_1 - \mu_2$ for Small Samples when $\sigma_1^2 = \sigma_2^2$.
Test Statistic: $\mathbf{t} = \dfrac{\bar{\mathbf{x}} - \bar{\mathbf{y}} - \delta_0}{\sqrt{\mathbf{s_p^2}\left(\frac{\mathbf{1}}{\mathbf{n_1}} + \frac{\mathbf{1}}{\mathbf{n_2}}\right)}}$, $s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{(n_1-1) + (n_2-1)}$.
($P(T_{n_1 + n_2 - 2} \geq t)$ is the area in the upper tail of the t-distribution greater than t.)

Testing Problem	Hypotheses	Reject H_0 if	P-Value
Upper One-sided	$H_0: \mu_1 - \mu_2 \leq \delta_0$ vs $H_1: \mu_1 - \mu_2 > \delta_0$	$t > t_{n_1+n_2-2,\alpha}$ $\Leftrightarrow$ $\bar{x} - \bar{y} > \delta_0 + t_{n_1+n_2-2,\alpha}\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$	$P(T_{n_1 + n_2 - 2} \geq t)$
Lower One-sided	$H_0: \mu_1 - \mu_2 \geq \delta_0$ vs $H_1: \mu_1 - \mu_2 < \delta_0$	$t < -t_{n_1+n_2-2,\alpha}$ $\Leftrightarrow$ $\bar{x} - \bar{y} < \delta_0 - t_{n_1+n_2-2,\alpha}\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$	$P(T_{n_1 + n_2 - 2} \leq t)$
Two-sided	$H_0: \mu_1 - \mu_2 = \delta_0$ vs $H_1: \mu_1 - \mu_2 \neq \delta_0$	$\lvert t\rvert > t_{n-1+n_2-2,\alpha/2}$ $\Leftrightarrow$ $\lvert\bar{x} - \bar{y} - \delta_0\rvert > t_{n_1+n_2-2,\alpha/2}\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$	$P(\lvert T_{n_1+n_2-2}\rvert \geq \lvert t\rvert)$ = $2P(\lvert T_{n_1+n_2-2}\rvert \geq \lvert t\rvert)$

The mean of the data is 69.4583. The variance is 466.899 and the standard deviation is 21.6078. We are interested in the null hypothesis $\mu = 80$. We can compute the t-statistic as follows.

TABLE 10. Level α Tests on $\mu_1 - \mu_2$ for Small Samples when $\sigma_1^2 \neq \sigma_2^2$.
Test Statistic: $\mathbf{t} = \dfrac{\bar{\mathbf{x}} - \bar{\mathbf{y}} - \delta_0}{\sqrt{\frac{\mathbf{s_1^2}}{\mathbf{n_1}} + \frac{\mathbf{s_2^2}}{\mathbf{n_2}}}}$, $w_1 = \dfrac{s_1^2}{n_1}$, $w_2 = \dfrac{s_2^2}{n_2}$, $\nu = \dfrac{(w_1 + w_2)^2}{\frac{w_1^2}{n_1 - 1} + \frac{w_2^2}{n_2 - 2}}$.
($P(T_\nu \geq t)$ is the area in the upper tail of the t-distribution greater than t.)

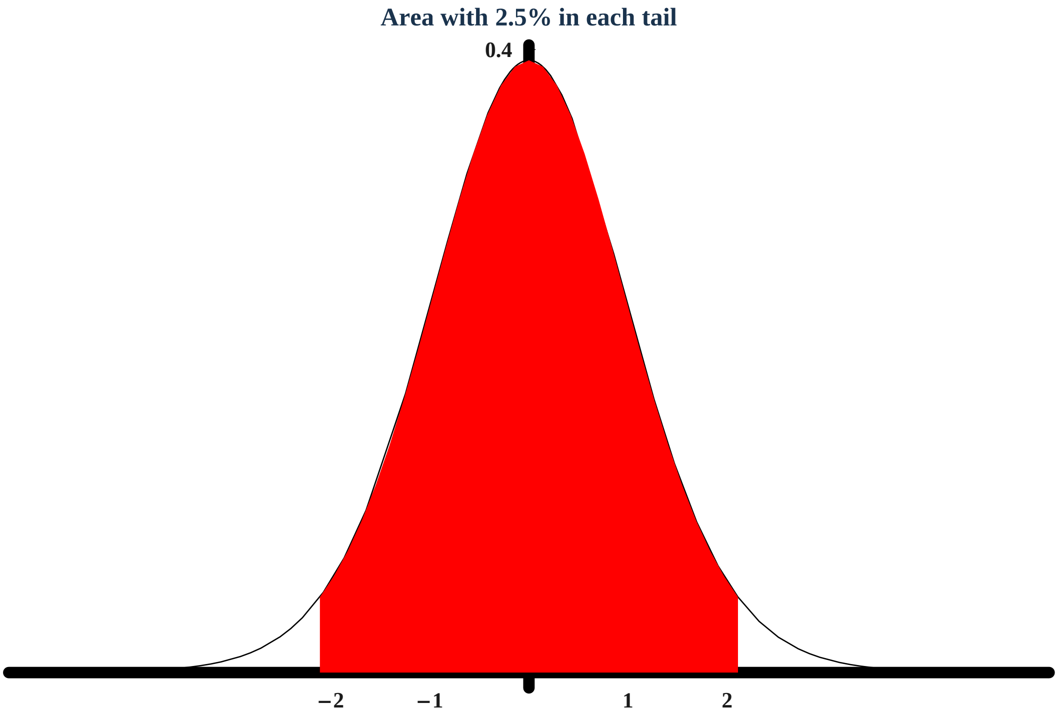
Testing Problem	Hypotheses	Reject H_0 if	P-Value
Upper One-sided	$H_0: \mu_1 - \mu_2 \leq \delta_0$ vs $H_1: \mu_1 - \mu_2 > \delta_0$	$t > t_{\nu,\alpha}$ $\Leftrightarrow$ $\bar{x} - \bar{y} > \delta_0 + t_{\nu,\alpha}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$	$P(T_\nu \geq t)$
Lower One-sided	$H_0: \mu_1 - \mu_2 \geq \delta_0$ vs $H_1: \mu_1 - \mu_2 < \delta_0$	$t < -t_{\nu,\alpha}$ $\Leftrightarrow$ $\bar{x} - \bar{y} < \delta_0 - t_{\nu,\alpha}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$	$P(T_\nu \leq t)$
Two-sided	$H_0: \mu_1 - \mu_2 = \delta_0$ vs $H_1: \mu_1 - \mu_2 \neq \delta_0$	$\lvert t \rvert > t_{\nu,\alpha/2}$ $\Leftrightarrow$ $\lvert \bar{x} - \bar{y} - \delta_0 \rvert > t_{\nu,\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$	$P(\lvert T_\nu \rvert \geq \lvert t \rvert)$ $=$ $2P(\lvert T_\nu \rvert \geq \lvert t \rvert)$

TABLE 11. Level α Tests for Equality of σ_1^2 and σ_2^2.
Test Statistic: $\mathbf{F} = \dfrac{\mathbf{S_1^2}}{\mathbf{S_2^2}}$.
($f_{n_1-1,\,n_2-1,\,\alpha}$ is the critical value such that α of the area of the F distribution with $n_1 - 1$ and $n_2 - 1$ degrees of freedom is greater than this value.)

Testing Problem	Hypotheses	Reject H_0 if
Upper One-sided	$H_0: \sigma_1^2 \leq \sigma_2^2$ vs $H_1: \sigma_1^2 > \sigma_2^2$	$F > f_{n_1-1,\,n_2-1,\,\alpha}$
Lower One-sided	$H_0: \sigma_1^2 \geq \sigma_2^2$ vs $H_1: \sigma_1^2 < \sigma_2^2$	$F < f_{n_1-1,\,n_2-1,\,1-\alpha}$
Two-sided	$H_0: \sigma_1^2 = \sigma_2^2$ vs $H_1: \sigma_1^2 \neq \sigma_2^2$	$F < f_{n_1-1,\,n_2-1,\,1-\alpha/2}$ or $F < f_{n_1-1,\,n_2-1,\,1-\alpha/2}$

$$
\begin{aligned}
t &= \frac{\bar{x} - \mu}{\frac{S}{\sqrt{n}}} \\
&= \frac{69.4583 - 80}{\frac{21.6078}{\sqrt{18}}} \\
&= -2.069\,83
\end{aligned}
\tag{72}
$$

Consider a one sided test that $\mu = 80$ with alternative $\mu < 80$. The value of the t statistic is $-2.069\,83$. We reject the null hypothesis if $-2.069\,83 < -t_{n-1,\,.05}$. From the t-tables $t_{17,\,.05} = 1.74$. Given that $-2.069\,83$ is less than -1.74, we reject the null hypothesis that $\mu = 80$.



To compute the p-value we find the probability that a value of the t-distribution lies to the left of -2.06983 or the right of 2.06983. Given that the t-distribution is symmetric we can find just one of these probabilities and multiply by 2. So what we need is the integral of the t-distribution from 2.06983 to ∞. This is given by

$$\int_{2.069\,83}^{\infty} \frac{\Gamma\left(\frac{r+1}{2}\right)}{\sqrt{\pi r}\,\Gamma\left(\frac{r}{2}\right)} \left(1 + \frac{w^2}{r}\right)^{\frac{-(r+1)}{2}} dw \tag{73}$$

where r is the degrees of freedom which in this case is 17. This integral can be evaluated numerically and will have value 0.0270083. If we double this we get 0.054 016 5. So 2.700 83% of the t-distribution with 17 degrees of freedom lies to the right of 2.069 83 and 2.700 83% of the distribution lies to the left of $-2.069\,83$.

11. More Hypothesis Testing Examples

Consider the income data for carpenters and house painters in from table 4 which we repeat here as table 12.

TABLE 12. Income Data for Carpenters and House Painters

	carpenters	painters
sample size	n_c = 12	n_p = 15
mean income	$\bar{c}$ = \$6000	$\bar{p}$ = \$5400
estimated variance	s_c^2 = \$565 000	s_p^2 = \$362 500

11.1. If it is known that $\sigma_x^2 = 600\,000$, test the hypothesis that $\mu_x = \$7000$ with $\alpha = .05$ against the alternative $\mu_x \neq \$7000$. The solution is as follows:

$$\begin{aligned} \bar{x} &\sim N\left(\mu_x, \frac{\sigma_x^2}{n_1}\right) \\ z &= \frac{\bar{x} - \mu_x}{\frac{\sigma_x}{\sqrt{n_1}}} \sim N(0,\, 1) \\ &= \frac{6000 - 7000}{\sqrt{\frac{600\,000}{12}}} = \frac{-1000}{223.61} = -4.47 \end{aligned} \tag{74}$$

The critical region is obtained by finding the value of γ such that the area in each tail of the distribution is equal to .025. This is done by consulting a normal table. The value of γ is 1.96. Thus if $|z| \geq 1.96$ we reject the hypothesis. Therefore, we reject $H_0: \ \mu_x = \$7000$.

11.2. Test the hypothesis that $\mu_x \geq \$7000$ versus the alternative $\mu < \$7000$ assuming that σ^2 is known as before.

We need a critical region such that only 5% of the distribution lies to the left of γ. Specifically we want a critical value for the test such that the probability of a type I error is equal to α. We write this as

$$\begin{aligned} P(\text{type I error}) &= P(\bar{x} < c \mid \mu = 7000) \\ &= P\left(z = \frac{\bar{x} - \mu_x}{\frac{\sigma_x}{\sqrt{n_1}}} < \frac{c - \mu_x}{\frac{\sigma_x}{\sqrt{n_1}}} \,\middle|\, \mu = 7000\right) \\ &= 0.05 \end{aligned} \tag{75}$$

Since the normal is symmetric we use the probability that 5% is the right to γ. This gives a critical value of 1.645.

Computing gives

$$\begin{aligned}\frac{c-7000}{\sqrt{\frac{600\,000}{12}}} &= -1.645\\ \Rightarrow c-7000 &= (-1.645)(\sqrt{50\,000})\\ \Rightarrow c-7000 &= -367.833\\ \Rightarrow c &= 6\,632.166\end{aligned} \tag{76}$$

So we reject the hypothesis if $\bar{x}$ is less than 6632.166. Alternatively this test can be expressed in terms of the standardized test statistic.

$$z = \frac{6000-7000}{\sqrt{\frac{600\,000}{12}}} = \frac{-1000}{223.61} = -4.47 \tag{77}$$

We reject the null hypothesis if $z \leq -1.645$. So we reject the null hypothesis.

11.3. Note that with two-tailed tests an equivalent procedure is to construct the $(1-\alpha)$ level confidence interval and reject the null hypothesis if the hypothesized value does not lie in that interval.

11.4. Test the same hypothesis in (11.1) for the case in which σ_x^2 is not known.
We use a t-test as follows.

$$\begin{aligned}\bar{x} &\sim N\left(\mu_x, \frac{\sigma_x^2}{n_1}\right)\\ \frac{\bar{x}-\mu_x}{s_{\bar{x}}} &= \frac{\bar{x}-\mu_x}{\frac{s_x}{\sqrt{n}}} \sim t(n-1)\\ &= \frac{6000-7000}{\sqrt{\frac{565\,000}{12}}} = \frac{-1000}{216.99}\\ &= -3.46\end{aligned} \tag{78}$$

The critical region is obtained by finding the value of γ such that the area in each tail of the distribution is equal to .025. This is done by consulting a t-table with $(n-1=11)$ degrees of freedom. The value of γ is 2.201. Thus if $|t| \geq 2.201$ we reject the hypothesis. Therefore, we reject $H_0: \ \mu_x = \$7000$.

11.5. Test the hypothesis that

$$\sigma_y^2 = 400\,000$$

against the alternative

$$\sigma_y^2 \neq 400\,000$$

We will use a two-tailed chi-square test. We must find levels of γ_1 and γ_2 such that the area in each tail is .025. If we consult a χ^2 table with 14 degrees of freedom we obtain 5.63 and 26.12. If the computed value is outside this range we reject the hypothesis. Computing gives

$$\frac{(n_2-1)s_y^2}{\sigma_y^2} \sim \chi^2(15-1)$$
$$= \frac{(14)(362\,500)}{400\,000} = 12.687 \tag{79}$$

Therefore fail to reject $H_0: \ \sigma_y^2 = 400\,000$.

11.6. Hypothesis Concerning the Equality of Two Means with $\sigma_1^2 = \sigma_2^2$. Let H_0 be $\mu_1 = \mu_2$ where we assume that $\sigma_1^2 = \sigma_2^2$. We reject the hypothesis if

$$|\bar{x} - \bar{y}| > t_{n_1+n_2-2,\alpha/2}\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \tag{80}$$

Assume that $\alpha = 0.05$. Remember that s_p^2 is given by

$$s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{(n_1-1) + (n_2-1)} \tag{81}$$

For the carpenter and housepainter data we have

$$\begin{aligned} n_c &= 12, & n_p &= 15 \\ \bar{c} &= 6000, & \bar{p} &= 5400 \\ \sigma_c^2 &= 600\,000, & \sigma_p^2 &= 565\,000 \end{aligned} \tag{82}$$

So s_p^2 is

$$\begin{aligned} s_p^2 &= \frac{(11)(600\,000) + (14)(565\,000)}{(11) + (14)} \\ &= \frac{14\,510\,000}{25} \\ &= 580\,400 \end{aligned} \tag{83}$$

The relevant t-value is $t_{25,0.025} = 2.06$. So we reject the hypothesis if

$$\begin{aligned} 600 &> 2.06\sqrt{580\,400\left(\frac{1}{12} + \frac{1}{15}\right)} \\ &> 2.06\sqrt{580\,400\left(\frac{5}{60} + \frac{4}{60}\right)} \\ &> 2.06\sqrt{580\,400\left(\frac{9}{60}\right)} \\ &> 2.06\sqrt{87\,060} \\ &> 607.82 \end{aligned} \tag{84}$$

So we do not reject the hypothesis that the means are equal. If we increased α to 0.01, we would reject the hypothesis. The test statistic is

$$\begin{aligned} t &= \frac{\bar{x} - \bar{y}}{\sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \\ &= \frac{600}{\sqrt{580\,400\left(\frac{1}{12} + \frac{1}{15}\right)}} \\ t &= \frac{\bar{x} - \bar{y}}{\sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \\ &= \frac{600}{\sqrt{87\,060}} \\ &= 2.033 \end{aligned} \tag{85}$$

If we look in a t-table with 25 degrees of freedom, we can see that the probability associated 2.033 is between 0.025 and 0.05 and very close to 0.025. Doubling this would give a p-value somewhere between 0.05 and 0.01 and. Specifically it is 0.0527.

11.7. Hypotheses Concerning Equality of Two Variances. Let H_0 be $\sigma_1^2 = \sigma_2^2$. If H_1 is $\sigma_1^2 > \sigma_2^2$, then the appropriate test statistic is the ratio of the sample variances since under the null hypothesis the variances are equal. We reject the hypothesis if

$$\frac{S_1^2}{S_2^2} > F_{\alpha,\, n_1 - 1,\, n_2 - 1} \tag{86}$$

If H_1 is $\sigma_1^2 < \sigma_2^2$ then the appropriate test statistic is the ratio of the sample variances since under the null hypothesis the variances are equal. We reject the hypothesis if

$$\frac{S_2^2}{S_1^2} > F_{\alpha,\, n_2 - 1,\, n_1 - 1} \tag{87}$$

If H_1 is $\sigma_1^2 \neq \sigma_2^2$ then the appropriate test statistic is the ratio of the sample variances since under the null hypothesis the variances are equal. We now have to decide which variance to put on top of the ratio. If the first is larger compute as in (86) and if the second is larger compute as in (87). Now however use $\alpha/2$ for significance. We reject the hypothesis if

$$\begin{aligned} &\frac{S_1^2}{S_2^2} > F_{\alpha/2,\, n_1 - 1,\, n_2 - 1} \\ \text{or} \quad &\frac{S_2^2}{S_1^2} > F_{\alpha/2,\, n_2 - 1,\, n_1 - 1} \end{aligned} \tag{88}$$

Consider this last case. Since s_x^2 is larger, put it on top in the ratio. The critical value of F is $F(11,\, 14 :\ .025) = 3.10$. For this case we obtain

$$\frac{s_x^2}{s_y^2} = \frac{565\,000}{362\,500} = 1.5586 \tag{89}$$

Since this is smaller than 3.10 we fail to reject $H_0 :\ \sigma_x^2 = \sigma_y^2$.