1. (Previously Problem 5 of PS 2) Consider the composite hypothesis testing problem:

\[ H_0 : \quad Y \text{ has density } p_0(y) = \frac{1}{2}e^{-|y|}, \quad y \in \mathcal{R} \]
\[ H_1 : \quad Y \text{ has density } p_0(y) = \frac{1}{2}e^{-|y-\theta|}, \quad y \in \mathcal{R}, \quad \theta > 0 \]

(a) Describe the locally most powerful \( \alpha \)-level test and derive its power function.

(b) Does a uniformly most powerful test exist? If so, find it and derive its power function. If not, find the generalized likelihood ratio test for \( H_0 \) versus \( H_1 \).

2. (Previously Problem 7 of PS 2) Consider the hypotheses of Problem 6 from PS 2 with \( \mu = \mu_1 > \mu_0 = 0 \) and \( \sigma^2 = \sigma_0^2 = \sigma_1^2 > 0 \). Does there exist a uniformly most powerful test of these hypotheses under the assumption that \( \mu \) is known and \( \sigma^2 \) is not? If so, find it and show that it is UMP. If not, show why and find the generalized likelihood ratio test.

3. Derivation of some important pdfs.

(a) Let \( Z_1, Z_2, \ldots, Z_n \) be i.i.d. and \( \mathcal{N}(0, 1) \). Then

\[ Y \overset{\text{def}}{=} \sum_{i=1}^{n} Z_i^2 \]

is said to be a chi-squared random variable with \( n \) degrees of freedom \( (\chi_n^2) \). Show that the pdf is

\[ f_Y(y) = \frac{e^{y/2}y^{(n/2)-1}}{2^{n/2}\Gamma(n/2)} \]

for \( y > 0 \) (and zero for \( y < 0 \)). Note \( \Gamma(\cdot) \) is the gamma function.

(b) Let \( X \sim \mathcal{N}(0, 1) \), \( Y \sim \chi_n^2 \) and \( X \) and \( Y \) statistically independent. Then we say (William Gosset, 1908)

\[ T = \frac{X}{\sqrt{Y/n}} \]

has “Student’s” \( T \) distribution with \( n \) degrees of freedom. Show that the pdf is

\[ f_T(t) = \frac{\Gamma((n+1)/2)}{\Gamma(n/2)\sqrt{n\pi}} \frac{1}{(1+t^2/n)^{(n+1)/2}} \]

for \( t \in \mathcal{R} \).

Hint: A workable approach to the solution involves using the pdf change of variable formula with the following invertible transformation from \((x, y)\): \( x \in \mathcal{R}, \ y > 0 \) to \((t, u)\): \( t \in \mathcal{R}, \ u > 0 \)

\[ t = \frac{x}{\sqrt{y/n}} \]
\[ u = \frac{y}{n} \]

to get the joint pdf of \( T \) and \( U \). Then one integrates over the variable \( U = u \) to get the marginal pdf for \( T \).

4. Sample mean, sample standard deviation, and related properties. Let \( X_1, X_2, \ldots \) be a sequence of real-valued random variables. The sample mean is defined to be

\[ \bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i \]
The sample variance is defined to be

\[ S_n^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2. \]

(a) Define deviations from the sample mean \( \bar{X}_n \) \( \overset{\text{def}}{=} X_i - \bar{X}_n, \) \( i = 1, 2, \ldots, n. \) Show that only \( n - 1 \) of the deviations can be picked independently. For this reason one says that \( S_n^2 \) has \( n - 1 \) degrees of freedom.

(b) Assume that the \( X_1, X_2, \ldots \) are i.i.d. with mean \( \mu \) and variance \( \sigma^2. \) Find the mean and variance of the sample mean \( \bar{X}_n. \)

(c) Find the mean of \( S_n^2. \)

(d) Assume that the \( X_1, X_2, \ldots \) are i.i.d. distributed as \( \mathcal{N}(\mu, \sigma^2). \) Here the goal is to characterize the joint distribution of \( \bar{X}_n \) and \( S_n^2. \)

i. Show that \( \bar{X}_n \sim \mathcal{N}(\mu, \sigma^2/n). \)

ii. Show that \( \bar{X}_n \) is statistically independent of the deviations \( \bar{X}_i, \) for \( i = 1, 2, \ldots, n. \)

\text{Hint: First establish the lemma:}

\text{Lemma: } X_1, X_2, \ldots, X_n \text{ i.i.d. of variance } \sigma^2. \text{ Then}

\[ \text{Cov}(\bar{X}_i, \bar{X}_n) = 0 \]

\text{for } i = 1, 2, \ldots, n.

\text{and use the fact that uncorrelated } \Rightarrow \text{ statistically independent for Gaussians.}

iii. Use the previous to show that \( \bar{X}_n \) is statistically independent of \( S_n^2. \)

iv. Finally show that

\[ \frac{(n-1)S_n^2}{\sigma^2} \sim \chi_{n-1}^2 \]

which completely characterizes the joint distribution of \( \bar{X}_n \) and \( S_n^2. \)
Consider the composite hypothesis testing problem:

\[ H_0 : \quad Y \text{ has density } p_0(y) = (1/2)e^{-|y|}, \ y \in \mathcal{R} \]
\[ H_1 : \quad Y \text{ has density } p_0(y) = (1/2)e^{-|y|-\theta}, \ y \in \mathcal{R}, \ \theta > 0 \]

(a) Describe the locally most powerful \( \alpha \)-level test and derive its power function.

(b) Does a uniformly most powerful test exist? If so, find it and derive its power function. If not, find the generalized likelihood ratio test for \( H_0 \) versus \( H_1 \).
The composite problem is

\[ H_0: \, P_0(y) = \frac{1}{2} e^{-|y|} \quad \text{for} \quad y \in \mathbb{R} \]

vs.

\[ H_1: \, P_0(y) = \frac{1}{2} e^{-|y-\theta|} \quad \text{for} \quad \theta > 0 \]

(a) From class we know that the locally most powerful test is of the form

\[
\delta_{\theta_0}^* (y) = \begin{cases} 
1 & \text{if } \frac{dP_0(y)}{d\theta} \bigg|_{\theta=0} = \eta P_0(y) \\
0 & \text{else}
\end{cases}
\]

so our first task is the computation of the derivative above. Note that

\[ P_0(y) = \begin{cases} 
\frac{1}{2} e^{-(y-\theta)} & y > \theta \\
\frac{1}{2} e^{-(y+\theta)} & y < \theta
\end{cases} \]

Technically we are not differentiable in \( \theta \) at \( \theta = y \) so we will stay away from this point.

\[ \theta > y \quad \frac{dP_0(y)}{d\theta} = -\frac{1}{2} e^{-(y+\theta)} \]

\[ \theta < y \quad \frac{dP_0(y)}{d\theta} = \frac{1}{2} e^{-(y-\theta)} \]

So for \( \theta = 0 \)

\[
\frac{dP_0(y)}{d\theta} \bigg|_{\theta=0} = \begin{cases} 
\frac{1}{2} e^{-y} & y > 0 \\
-\frac{1}{2} e^{y} & y < 0
\end{cases}
\]

\[ = \text{sgn}(y) \frac{1}{2} e^{-|y|} \quad y \neq 0 \]

\[ \therefore \quad \frac{\frac{dP_0(y)}{d\theta}}{P_0(y)} \bigg|_{\theta=0} = \text{sgn}(y) \]
Therefore the LMP test is of the form

\[ \tilde{\delta}_{E_0}(y) = \begin{cases} 
1 & \text{if } \text{sgn}(y) \geq \eta \\
0 & \text{if } \text{sgn}(y) < \eta 
\end{cases} \]

To set the threshold for \( \alpha \)-level testing we must consider the probability \( P_0 \{ \text{sgn}(y) > \eta \} \) which is 1 minus the distribution of the statistic \( \text{sgn}(y) \). Note that

\[ P_0 \{ \text{sgn}(y) = 1 \} = \frac{1}{2} = P_0 \{ \text{sgn}(y) = -1 \} \]

Therefore the distribution function is of the form:

\[ P_0 \{ \text{sgn}(y) \leq \eta \} \]

\[ P_0 \{ \text{sgn}(y) > \eta \} \]

Clearly if \( 0 \leq \alpha < \frac{1}{2} \) the smallest \( \eta \) for which \( P_0 \{ \text{sgn}(y) > \eta \} \leq \alpha \) is \( \eta = 1 \) and if \( \frac{1}{2} \leq \alpha < 1 \) the smallest \( \eta \) is \( \eta = -1 \). So pick

\[ \eta = \begin{cases} 
1 & \text{if } 0 \leq \alpha < \frac{1}{2} \\
-1 & \text{if } \frac{1}{2} \leq \alpha < 1 
\end{cases} \]
Randomization is required here. From our proof of the NP lemma, the randomization is:

\[ \Gamma = \frac{\alpha - P_{0}\{ \text{sgn}(Y) > \eta \}}{P_{0}\{ \text{sgn}(Y) = \eta \}} = \begin{cases} 2\alpha - 1 & \frac{1}{2} \leq \alpha < 1 \\ 2\alpha & 0 \leq \alpha < \frac{1}{2} \end{cases} \]

Then for \(0 \leq \alpha < \frac{1}{2}\) the test is:

\[ \tilde{\delta}_{\tilde{\Theta}},(y) = \begin{cases} \frac{1}{2\alpha} & \text{if } \text{sgn}(y) = +1 \\ 0 & \text{else} \end{cases} \]

and if \(\frac{1}{2} \leq \alpha < 1\) the test is:

\[ \tilde{\delta}_{\tilde{\Theta}},(y) = \begin{cases} \frac{1}{2\alpha - 1} & \text{if } \text{sgn}(y) = -1 \\ 0 & \text{else} \end{cases} \]

These tests can be simplified a bit as:

\[ \tilde{\delta}_{\tilde{\Theta}},(y) = \begin{cases} 2\alpha & y > 0 \\ 0 & y \leq 0 \end{cases} \]

\[ \tilde{\delta}_{\tilde{\Theta}},(y) = \begin{cases} \frac{1}{2\alpha - 1} & y < 0 \\ 0 & y \geq 0 \end{cases} \]

\[ \frac{1}{2} \leq \alpha < 1 \]

The next task is to compute the detection probability for these tests. For fixed \(\Theta > 0\):

\[ P_{D}(\tilde{\delta}_{\tilde{\Theta}}; \Theta) = P_{\Theta}\{ \text{sgn}(Y) > \eta \} + \frac{1}{2} P_{\Theta}\{ \text{sgn}(Y) = \eta \} \]

It will be useful to consider the two cases above separately.
For $0 \leq \alpha < \frac{1}{2}$

$$P_D(\tilde{S}_{20}; \theta) = 2\alpha P_{\theta}\{Y > 0\} = 2\alpha \int_{0}^{\infty} \frac{1}{2} e^{-\frac{1}{2}y-\theta} dy$$

$$= 2\alpha \left\{ \int_{0}^{\infty} \frac{1}{2} e^{-(y-\theta)} dy + \int_{\theta}^{\infty} \frac{1}{2} e^{-(y-\theta)} dy \right\}$$

$$= 2\alpha \left\{ \frac{1}{2} e^{\theta} \right\}_{y=0}^{y=\infty} - \frac{1}{2} e^{-(y-\theta)} \right\}_{y=\theta}^{y=\infty}\right\}$$

$$= 2\alpha \left\{ \frac{1}{2} - \frac{1}{2} e^{-\theta} + \frac{1}{2} \right\} = \alpha (2 - e^{-\theta})$$

For $\frac{1}{2} \leq \alpha < 1$

$$P_D(\tilde{S}_{20}; \theta) = P_{\theta}\{Y > 0\} + (2\alpha - 1) P_{\theta}\{Y < 0\}$$

$$= \int_{0}^{\infty} \frac{1}{2} e^{-\frac{1}{2}y-\theta} dy + (2\alpha - 1) \left[ \int_{0}^{\infty} \frac{1}{2} e^{-\frac{1}{2}y-\theta} dy \right]$$

$$= (1 - \frac{1}{2} e^{-\theta}) + (2\alpha - 1) \left[ \frac{1}{2} e^{y-\theta} \right]_{y=0}^{y=\infty}$$

$$= (1 - \frac{1}{2} e^{-\theta}) + (2\alpha - 1) \frac{1}{2} e^{-\theta}$$

$$= 1 - \frac{1}{2} e^{-\theta} + \alpha e^{-\theta} - \frac{1}{2} e^{-\theta} = 1 + (\alpha - 1) e^{-\theta}$$

(b) To decide the existence of a UMP test we need to go back to the LRT and find the form of the critical regions. The likelihood ratio for a particular value of $\theta$

$$L_{\theta}(y) = \frac{\frac{1}{2} e^{-\frac{1}{2}y-\theta}}{\frac{1}{2} e^{-\frac{1}{2}y}} \quad ye \mathbb{R} \quad \theta > 0.$$

$$= e^{-\frac{1}{2}y-\theta+\frac{1}{2}y} = e^{-\frac{1}{2}y-\theta}.$$
Since the logarithm is monotone increasing, NP tests compare the statistic $|y_1 - y - \Theta|$ to a threshold. That is, the NP test for a fixed $\Theta$ is of the form

$$
\delta_\Theta(y) = \begin{cases} 
1 & \text{if } |y_1 - y - \Theta| > \eta \\
0 & \text{otherwise}
\end{cases}
$$

meaning that the critical region of the test is

$$
\Pi_\Theta = \left\{ y \in \mathbb{R} : |y_1 - y - \Theta| > \eta \right\}.
$$

To get a better feel for this critical region plot the test statistic $T_\Theta(y) = |y_1 - y - \Theta|$ as a function of $y$.

Hence if $\eta < -\Theta$ we see that $\Pi_\Theta = \mathbb{R}$ while if $\eta > \Theta$ one has that $\Pi_\Theta = \emptyset$. If $-\Theta \leq \eta < \Theta$ it is easy to see that $\Pi_\Theta = \left( \frac{\eta + \Theta}{2}, \infty \right)$.

The process of setting the threshold $\eta$ starts with the computation of $P_\Theta(\Pi_\Theta)$. Then

$$
\eta < -\Theta \quad \Rightarrow \quad \Pi_\Theta = \mathbb{R} \quad \Rightarrow \quad P_\Theta(\Pi_\Theta) = 1
$$

$$
\eta > \Theta \quad \Rightarrow \quad \Pi_\Theta = \emptyset \quad \Rightarrow \quad P_\Theta(\Pi_\Theta) = 0
$$
If \(-\theta \leq \eta < \theta\) we find
\[
P_0(\eta_0) = \int_{\eta + \theta \over 2}^{\infty} \frac{1}{2} e^{-y} dy = \frac{1}{2} \left[ \int_{\eta + \theta \over 2}^{\infty} e^{-y} dy - \frac{1}{2} e^{-y} \right]_{\eta + \theta \over 2}^{\infty} = \frac{1}{2} e^{-{\eta + \theta \over 2}}
\]
where we note that \(\eta + \theta \geq 0\) for the case at hand. Summarizing our finding
\[
P_0(\eta_0) = \begin{cases} 
1 & \eta < -\theta \\
\frac{1}{2} e^{-{\eta + \theta \over 2}} & -\theta \leq \eta < \theta \\
0 & \eta \geq \theta 
\end{cases}
\]

We see from this that given a desired level \(\alpha\), we must know \(\theta\) in order to set the threshold. Thus it would seem that UMP tests do not exist. But see the following pages.

The generalized likelihood ratio test uses the statistic
\[
\sup_{\theta > 0} \frac{P_0(y)}{P_0(y)} = \sup_{\theta > 0} \frac{1}{2} e^{-\eta - \theta} = \frac{1}{2} e^{-\eta} = \sup_{\theta > 0} \exp\{\chi \eta - \chi \eta - \Theta\}
\]
and since \(\exp\{\cdot\}\) is monotone increasing in its argument the statistic for the generalized LRT is
\[
\exp\left\{\sup_{\Theta > 0} [\chi \eta - \chi \eta - \Theta]\right\} = \begin{cases} 
e^{y} & y \geq 0 \\
1 & y < 0
\end{cases}
\]
Performance of Tests

Consider first the optimum tests assuming $\Theta > 0$ is known. We showed that critical regions were of the form:

$$\Gamma_\Theta = \{ \gamma \in \mathbb{R} : \left| y_1 - y_2 - \Theta \right| > \eta \}$$

We also found:

$$P_\Theta(\Gamma_\Theta) = P_\Theta(\left| y_1 - y_2 - \Theta \right| > \eta) = \begin{cases} \frac{1}{2} e^{-(\eta+\Theta)^2} & \eta < -\Theta \\ \frac{1}{2} e^{-(\eta+\Theta)^2} & -\Theta \leq \eta < \Theta \\ 0 & \eta \geq \Theta \end{cases}$$

For setting the threshold there are several cases to consider (we only worry about $\alpha$ strictly between 0 and 1).

**Case:** $\frac{1}{2} < \alpha < 1$

Then $\eta = -\Theta$ and the randomization is

$$\gamma = \frac{\alpha - \frac{1}{2}}{\sqrt{2}} = 2\alpha - 1$$

**Case:** $\frac{1}{2} e^{-\Theta} < \alpha \leq \frac{1}{2}$

Obtain $\eta$ by solving $\frac{1}{2} e^{-(\eta+\Theta)^2} = \alpha \Rightarrow \eta = -\Theta - 2 \log(2\alpha)$

Here no randomization is required.

**Case:** $0 < \alpha \leq \frac{1}{2} e^{-\Theta}$

Then $\eta = \Theta$ and the randomization is

$$\gamma = \frac{\alpha}{\frac{1}{2} e^{-\Theta}} = 2\alpha e^{\Theta}$$
Computation of Performance

Case $\frac{1}{2} < \alpha < 1 \Rightarrow \eta = -\Theta$, $Y = 2\alpha - 1$

$$\tilde{\delta}_\Theta(y) = \begin{cases} \frac{1}{2} & \text{if } |y| - 1_y - \Theta < -\Theta \\ 2\alpha - 1 & \text{if } |y| - 1_y - \Theta > -\Theta \end{cases}$$

(Look at plot of $T_\Theta(y)$.

$$= \begin{cases} \frac{1}{2} & \text{if } y > 0 \\ 2\alpha - 1 & \text{if } y \leq 0 \end{cases}$$

The performance is

$$E_\Theta\{\tilde{\delta}_\Theta(y)\} = P_\Theta\{Y > 0\} + (2\alpha - 1) P_\Theta\{Y \leq 0\}$$

for this we need to again look at the density $P_\Theta$.

$$P_\Theta(y) = \frac{1}{2} e^{-|y| - \Theta} \quad y \in \mathbb{R}$$

$$P_\Theta\{Y \leq 0\} = \int_{-\infty}^{0} \frac{1}{2} e^{-y - \Theta} \, dy = \frac{1}{2} e^{-\Theta} \left[ e^y \right]_{-\infty}^{0} = \frac{1}{2} e^{-\Theta}$$

$$P_\Theta\{Y > 0\} = \frac{1}{2} + \frac{1}{2} - \frac{1}{2} e^{-\Theta} = 1 - \frac{1}{2} e^{-\Theta}$$

$$\therefore E_\Theta\{\tilde{\delta}_\Theta(y)\} = (1 - \frac{1}{2} e^{-\Theta}) + (2\alpha - 1) \frac{1}{2} e^{-\Theta}$$

$$= 1 + (\alpha - 1) e^{-\Theta}$$

Case $\frac{1}{2} e^{-\Theta} < \alpha < \frac{1}{2} \Rightarrow \eta = -\Theta - 2\log(2\alpha) \doteq \eta(\alpha)$

$$\tilde{\delta}_\Theta(y) = \begin{cases} \frac{1}{2} & \text{if } |y| - 1_y - \Theta < -\Theta - 2\log(2\alpha) \\ 0 & \text{if } |y| - 1_y - \Theta > -\Theta - 2\log(2\alpha) \end{cases}$$

Note that over the range of $\alpha$ above $\eta(\alpha)$ decreases monotonically from $\Theta$ to $-\Theta$. Thus the intersection

$$z_y - \Theta = \eta(\alpha)$$
\[ 2y - \theta = -\theta - 2\log(2x) \]
\[ y = -\log(2x). \]

An equivalent test is:

\[ \tilde{S}_\theta(y) = \begin{cases} 
\frac{1}{2} & \text{if } y > -\log(2x) \\
0 & \text{if } y < -\log(2x) 
\end{cases} \]

\[ E_\theta\{\tilde{S}_\theta(y)\} = P_\theta\{Y > -\log(2x)\} \]

\[ \frac{-\log(2x)}{2} + \frac{1}{2} - \int_{-\infty}^{-\log(2x)} \frac{1}{2} e^{y-\theta} dy = 1 - \frac{1}{2} e^{-\theta} - \log(2x) \]

For the range of \( \alpha \)'s considered, this goes from \( \theta \) to 0.

Case \( 0 < \alpha \leq \frac{1}{2} e^{-\theta} \Rightarrow \gamma = \theta, \quad y = zae^\theta \)

\[ \tilde{S}_\theta(y) = \begin{cases} 
\frac{1}{2} & \text{if } |y - \gamma| > \theta \\
0 & \text{if } |y - \gamma| \leq \theta 
\end{cases} \]

\[ = \begin{cases} 
zae^\theta 
& \text{if } y > \theta \\
0 
& \text{if } y < \theta 
\end{cases} \]

\[ E_\theta\{\tilde{S}_\theta(y)\} = zae^\theta P_\theta\{Y > \theta\} = \alpha e^\theta \]
Summarizing the Cases

\[
\begin{array}{ccc}
\frac{1}{2} < \alpha < 1 & \frac{1}{2} e^\Theta < \alpha \leq \frac{1}{2} & 0 < \alpha \leq \frac{1}{2} e^\Theta \\
\eta & -\Theta & -\Theta - 2\log(2\alpha) & \Theta \\
\gamma & 2\alpha - 1 & 0 & 2\alpha e^\Theta \\
P_0 & 1 + (\alpha - 1) e^\Theta & 1 - \frac{e^{-\Theta}}{4\alpha} & \alpha e^\Theta \\
\end{array}
\]
Generalized LRT

Found \( G(y) = \begin{cases} e^y & y > 0 \\ 1 & y < 0 \end{cases} \)

To set the threshold we need to compute \( P_0(G(y) > \eta) \) as a function of \( \eta \).

Case \( \eta < 1 \quad G(y) > \eta \quad \forall y \in \mathbb{R} \)

Case \( \eta \geq 1 \quad G(y) > \eta \iff y > \log \eta \)

Recall the pdf under hypothesis 0. Then for \( \eta \geq 1 \):

\[
P_0 \{ Y > \log \eta \} = \int_{y=\log \eta}^{\infty} \frac{1}{2} e^{-y} dy = -\frac{1}{2} e^{-y} \bigg|_{\log \eta}^{\infty} = \frac{1}{2} e^{-\log \eta} = \frac{1}{2} e^{-\log \eta} = \frac{1}{2} \eta\]

\[
P_0(G(Y) > \eta) \quad \log \eta = \frac{1}{2} e^{-\log \eta} = \frac{1}{2} \eta\]

So for \( 0 < \alpha \leq \frac{1}{2} \), no randomization is needed and the threshold is obtained by solving

\[
\frac{1}{2} \eta = \alpha \quad \Rightarrow \quad \eta = \frac{1}{2} \alpha
\]
For such \( \alpha \) the test is

\[
\tilde{\delta}_{GL}(y) = \begin{cases} 
1 & G(y) > \eta = \frac{1}{2} \alpha \\
0 & \text{others}
\end{cases}
\]

\[
= \begin{cases} 
1 & y > \log \eta = -\log 2\alpha \\
0 & \text{others}
\end{cases}
\]

Then

\[
E_\Theta \{ \tilde{\delta}_{GL}(Y) \} = P_\Theta \{ Y > -\log 2\alpha \}
\]

Two cases to consider depending on if \(-\log 2\alpha\) is greater or less than \(\Theta\). In the breakpoint \(-\log 2\alpha = \Theta \Rightarrow \alpha = \frac{1}{2} e^{-\Theta}\) (Note this is same breakpoint as we had before.)

Thus cases are \(0 < \alpha < \frac{1}{2} e^{-\Theta}\) and \(\frac{1}{2} e^{-\Theta} < \alpha < \frac{1}{2}\)

Case: \(\frac{1}{2} e^{-\Theta} < \alpha \leq \frac{1}{2}\)

Same as before. We found

\[
E_\Theta \{ \tilde{\delta}_{GL}(Y) \} = 1 - \frac{e^{-\Theta}}{4-\alpha}
\]

Case: \(0 < \alpha < \frac{1}{2} e^{-\Theta}\)

\[
E_\Theta \{ \tilde{\delta}_{GL}(Y) \} = \int_{-\log 2\alpha}^{\infty} \frac{1}{2} e^{-y+\Theta} dy = \frac{1}{2} e^\Theta \left( -e^{-y} \right) \bigg|_{-\log 2\alpha}^{\infty}
\]

\[
= -\frac{1}{2} e^\Theta \left[ 0 - e^{\log 2\alpha} \right] = \frac{2\alpha}{2} e^\Theta = \alpha e^\Theta
\]
Consider the case $\frac{1}{2} < \alpha < 1$

For this range of $\alpha$ the threshold is $\eta = 1$ and the randomization is

$$\gamma = \frac{\alpha - \frac{1}{2}}{\frac{1}{2}} = 2\alpha - 1$$

The test is

$$\delta_{GL}(y) = \begin{cases} 1 & y > 2\alpha - 1 \\ G(y) = \frac{\eta}{2} & 2\alpha - 1 > y \geq 0 \\ 0 & y \leq 0 \end{cases}$$

Of course, this has the same performance.
Consider the hypotheses of the previous problem with $\mu = \mu_1 > \mu_0 = 0$ and $\sigma^2 = \sigma_0^2 = \sigma_1^2 > 0$. Does there exist a uniformly most powerful test of these hypotheses under the assumption that $\mu$ is known and $\sigma^2$ is not? If so, find it and show that it is UMP. If not, show why and find the generalized likelihood ratio test.

Consider the following pair of hypotheses concerning a sequence $Y_1, Y_2, \ldots, Y_n$ of random variables

- $H_0: \quad Y_k \sim \mathcal{N}(\mu_0, \sigma_0^2), \quad k = 1, 2, \ldots, n$
- $H_1: \quad Y_k \sim \mathcal{N}(\mu_1, \sigma_1^2), \quad k = 1, 2, \ldots, n$

where $\mu_0, \mu_1, \sigma_0^2$, and $\sigma_1^2$ are known constants.

(a) Show that the likelihood ratio can be expressed as a function of the parameters $\mu_0, \mu_1$, $\sigma_0^2$, and $\sigma_1^2$, and the quantities $\sum_{k=1}^{n} Y_k^2$ and $\sum_{k=1}^{n} Y_k$.

(b) Describe the Neyman–Pearson test for the two cases ($\mu_0 = \mu_1, \sigma_1^2 > \sigma_0^2$) and ($\sigma_0^2 = \sigma_1^2, \mu_1 > \mu_0$).

(c) Find the threshold and ROCs for the case $\mu_0 = \mu_1, \sigma_1^2 > \sigma_0^2$ with $n = 1$. 
\[ \mu = \mu_1, \mu_0 = 0 \quad \sigma^2 = \sigma_0^2 = \sigma_1^2 > 0 \]

\[ H_0: Y_k \sim \eta(0, \sigma^2) \quad k = 1, 2, \ldots, n \quad (\mu \text{ is known}) \]

\[ H_1: Y_k \sim \eta(\mu, \sigma^2) \]

Here both hypotheses are composite, because \( \sigma^2 \) is not known. The entire parameter space for the problem is

\[ \Lambda = \{(0, \mu) \times \mathbb{R}^+ \text{ with } \mu \text{ is fixed, } \mathbb{R}^+ \text{ is pos. real numbers.} \}

\[ \Lambda = \Lambda_0 \cup \Lambda_1 \quad \text{(a disjoint union)} \]

\[ \Lambda_0 \text{ set of ordered pairs } (0, \sigma^2) \quad \sigma^2 > 0 \]

\[ \Lambda_1 \text{ set of ordered pairs } (\mu, \sigma^2) \quad \sigma^2 > 0 \]

In this case a UMP test of level \( \alpha \) (if exists) is one that maximizes

\[ P_0(\hat{\sigma}^2; \mu, \sigma^2) \quad \text{for every } \sigma^2 > 0 \]

subject to

\[ P_F(\hat{\sigma}^2; 0, \sigma^2) \leq \alpha \quad \text{for every } \sigma^2 > 0. \]

Another way to interpret the problem is to hypothesize the existence of iid random variables \( N_k, k = 1, 2, \ldots, n \) where each is distributed \( \eta(0, \sigma^2) \). Then we have

\[ H_0: Y_k = N_k \quad k = 1, 2, \ldots, n \]

\[ H_1: Y_k = \mu + N_k \]

This would be the problem of detecting a known signal in noise. There is a slight distinction between this and the first composite problem. Here though \( \sigma^2 \) is unknown, it is the same in either hypothesis.
Aside: The level $\alpha$ test

Specializing Prob. 2 to the current situation the NP test is of the form

$$\delta_{NP}(y) = \begin{cases} 
  1 & \text{if } \sum_{k=1}^{n} y_k > \frac{z'}{\sqrt{n}} \left[ \log y + n \frac{\mu^2}{2\sigma^2} \right] \\
  0 & \text{otherwise} 
\end{cases}$$

Of course, we really need to know how to set $z'$ for level $\alpha$. Note that

$$T(Y) = \sum_{k=1}^{n} Y_k$$

is a linear comb. of rand. variables which are iid and Gaussian under either hypothesis. Thus $T(Y)$ is a Gaussian random variable.

Under $H_0$

$$E_{0}\{T(Y)\} = 0 \quad Var_{0}\{T(Y)\} = n\sigma^2$$

$$\therefore T(Y) \sim \mathcal{N}(0, n\sigma^2) \quad \text{under } H_0$$

Randomization will not be necessary. To set the threshold we must solve

$$P_{0}\{T(Y) > z'\} = 1 - P_{0}\{T(Y) \leq z'\} = 1 - P_{0}\left\{ \frac{T(Y)}{\sqrt{n}\sigma} \leq \frac{z'}{\sqrt{n}\sigma} \right\}$$

$$= 1 - \Phi\left( \frac{z'}{\sqrt{n}\sigma} \right) = \alpha$$

$$\Rightarrow z' = \sqrt{n}\sigma \Phi^{-1}(1-\alpha)$$

Under $H_1$

$$E_1\{T(Y)\} = n\mu \quad Var_1\{T(Y)\} = n\sigma^2$$

$$\therefore T(Y) \sim \mathcal{N}(n\mu, n\sigma^2) \quad \text{under } H_1$$
\[ P_D(\delta_{NP}) = P_1(\tau(y) > z') = 1 - P_1\{ \tau(y) \leq z' \} \]
\[ = 1 - P_1\left\{ \frac{\tau(y) - \eta \mu}{\sqrt{n} \sigma} \leq \frac{z' - \eta \mu}{\sqrt{n} \sigma} \right\} \]
\[ = 1 - \Phi\left( \frac{z' - \eta \mu}{\sqrt{n} \sigma} \right) \]
\[ = 1 - \Phi\left( \frac{\sqrt{n} \sigma \Phi^{-1}(1-\alpha) - \eta \mu}{\sqrt{n} \sigma} \right) \]
\[ = 1 - \Phi\left( \Phi^{-1}(1-\alpha) - \frac{\mu}{\sqrt{n}} \right) \]

I've actually done these computations assuming that \( \sigma \) is the same under both hypotheses. From the threshold computation the critical region is:
\[ \Pi_{\sigma^2} = \left\{ y \in \mathbb{R}^n : \sum_{k=1}^{n} y_k > \sqrt{n} \sigma \Phi^{-1}(1-\alpha) \right\} \]
which certainly depends upon \( \sigma \). Therefore, a UMP test does not exist.

The Generalized LRT

From the iid assumption
\[ p_1(y) = (2\pi)^{n/2} \sigma^n \exp\left\{ -\frac{1}{2\sigma^2} \sum_{k=1}^{n} (y_k - \mu)^2 \right\} \]
\[ p_0(y) = (2\pi)^{n/2} \sigma^n \exp\left\{ -\frac{1}{2\sigma^2} \sum_{k=1}^{n} y_k^2 \right\} \]
\[ G(y) = \frac{\sup_{\sigma^2 > 0} p_1(y)}{\sup_{\sigma^2 > 0} p_0(y)} \]
Consider \( y \) as fixed and find value of \( \sigma^2 \) maximizing \( p_1(y) \).

Have

\[
\frac{d}{d\sigma^2} (p_1(y)) = -n \left( \frac{2\pi}{\sigma^2} \right)^{-\frac{n}{2}} \sigma^{-\frac{n}{2}} \exp\left\{ -\frac{n}{2\sigma^2} \sum_{k=1}^{n} (y_k - \mu)^2 \right\} \exp\left\{ -\frac{n}{2\sigma^2} \sigma^2 \right\} \left( \sigma^2 \right)^{-\frac{n}{2}} \exp\left\{ -\frac{n}{2\sigma^2} \right\} (\sigma^2 \sum_{k=1}^{n} (y_k - \mu)^2) \\
= 0
\]

\[
-n \sum_{k=1}^{n} (y_k - \mu)^2 + \frac{1}{\sigma^3} \sum_{k=1}^{n} (y_k - \mu)^2 = 0
\]

\[
-n \sigma^2 + \sum_{k=1}^{n} (y_k - \mu)^2 = 0 \quad \Rightarrow \quad \sigma^2 = \frac{1}{n} \sum_{k=1}^{n} (y_k - \mu)^2
\]

Similarly, setting \( \frac{d}{d\sigma^2} p_0(y) = 0 \) get \( \sigma^2 = \frac{1}{n} \sum_{k=1}^{n} y_k^2 \).

We plug these estimates back in to find

\[
\sup_{\sigma^2 > 0} p_1(y) = (2\pi)^{-n/2} \left( \frac{1}{n} \sum_{k=1}^{n} (y_k - \mu)^2 \right)^{-n/2} \exp\left\{ -\frac{n}{2} \right\} \left( \frac{1}{n} \sum_{k=1}^{n} (y_k - \mu)^2 \right)^{-n/2} \exp\left\{ -\frac{n}{2} \right\}
\]

\[
\sup_{\sigma^2 > 0} p_0(y) = (2\pi)^{-n/2} \left( \frac{1}{n} \sum_{k=1}^{n} y_k^2 \right)^{-n/2} \exp\left\{ -\frac{n}{2} \right\} \left( \frac{1}{n} \sum_{k=1}^{n} y_k^2 \right)^{-n/2} \exp\left\{ -\frac{n}{2} \right\}
\]

\[
\therefore \quad G(y) = \left( \frac{\sum_{k=1}^{n} y_k^2}{\sum_{k=1}^{n} (y_k - \mu)^2} \right)^{n/2}
\]

Or could just use the statistic

\[
\hat{G}(y) = \frac{\sum_{k=1}^{n} y_k^2}{\sum_{k=1}^{n} (y_k - \mu)^2}
\]
Derivation of some important pdfs.

(a) Let $Z_1, Z_2, \ldots, Z_n$ be i.i.d. and $\mathcal{N}(0, 1)$. Then

$$Y \overset{\text{def}}{=} \sum_{i=1}^{n} Z_i^2$$

is said to be a chi-squared random variable with $n$ degrees of freedom ($\chi_n^2$). Show that

$$f_Y(y) = \frac{e^{y/2} y^{(n/2)-1}}{2^{n/2} \Gamma(n/2)}$$

for $y > 0$ (and zero for $y < 0$). Note $\Gamma(\cdot)$ is the gamma function.

(b) Let $X \sim \mathcal{N}(0, 1)$, $Y \sim \chi_n^2$ and $X$ and $Y$ statistically independent. Then we say (William Gosset, 1908)

$$T = \frac{X}{\sqrt{Y/n}}$$

has "Student's" $T$ distribution with $n$ degrees of freedom. Show that the pdf is

$$f_T(t) = \frac{\Gamma((n+1)/2)}{\Gamma(n/2)\sqrt{n}} \frac{1}{(1 + t^2/n)^{(n+1)/2}}$$

for $t \in \mathcal{R}$.

*Hint:* A workable approach to the solution involves using the pdf change of variable formula with the following invertible transformation from $(x, y)$: $x \in \mathcal{R}$, $y > 0$ to $(t, u)$: $t \in \mathcal{R}$, $u > 0$

$$t = \frac{x}{\sqrt{y/n}}$$

$$u = y$$

to get the joint pdf of $T$ and $U$. Then one integrates over the variable $U = u$ to get the marginal pdf for $T$. 
Re: \( X^2, T \) and \( F \) Distributions

If \( Z_1, Z_2, \ldots, Z_n \) are iid and \( N(0,1) \) then

\[
Y \sim \sum_{i=1}^{n} Z_i^2
\]

is said to be a chi-squared (\( X^2 \)) random variable with \( n \) degrees of freedom.

It is fairly easy to derive the pdf [Ross 1998 p 267]

\[
f_Y(y) = \frac{\Gamma(n/2)}{2^{n/2} (\Gamma(n/2))} y^{(n/2)-1} \quad y > 0 \quad (\text{and zero for} y < 0)
\]

\( \Gamma \) function

"Student's" \( T \) Distribution [William Gosset, 1908]

Let \( X \sim N(0,1) \) and \( Y \sim X_N^2 \) and \( X \perp Y \). Then

\[
T = \frac{X}{\sqrt{Y/N}}
\]

has a \( T \) distribution. Said to have \( N \) degrees of freedom.

The pdf for \( T \) is (n degrees of freedom)

\[
f_T(t) = \frac{1}{\Gamma(n/2) \sqrt{n\pi}} \frac{\Gamma((n+1)/2)}{(1 + t^2/n)^((n+1)/2)} \quad t \in \mathbb{R}
\]
Derivation of Tpdf \([\text{Scharf p 175}]\)

Define a transformation \((t, u) = W(x, y)\) by

\[
t = \frac{x}{\sqrt{y/n}}, \quad x \in \mathbb{R}, y > 0
\]

\[
u = y
\]

This transformation is invertible and can write \((x, y) = W^{-1}(t, u)\) where

\[
x = t \sqrt{u/n}
\]

\[
y = u
\]

and the Jacobian of the transformation \(W^{-1}\) is

\[
|J| = \det \begin{bmatrix}
x_x & x_u \\
y_x & y_u
\end{bmatrix} = \det \begin{bmatrix}
\frac{\partial x}{\partial t} & \frac{\partial x}{\partial u} \\
\frac{\partial y}{\partial t} & \frac{\partial y}{\partial u}
\end{bmatrix} = \det \begin{bmatrix}
\sqrt{u/n} & \frac{1}{2} t (u/n)^{-1/2} \\
0 & 1
\end{bmatrix}
\]

\[
= \sqrt{u/n}
\]

Then, the C.O.V. formula is:

\[
f_{T,U}(t,u) = f_{X,Y}(t \sqrt{u/n}, u) |J|
\]

where the joint pdf \(f_{X,Y}\) has product form since \(X \perp Y\) from our hypothesis:

\[
f_{X,Y}(x, y) = f_X(x) f_Y(y)
\]

\[
= \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \frac{1}{\Gamma(n/2)} \frac{1}{2^{n/2}} y^{(n/2) - 1} e^{-y/2}, \quad x \in \mathbb{R}, y > 0
\]
Substituting we get the joint distribution of $T, U$:

$$f_{T,U}(t,u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2n}} \frac{e^{-\frac{t^2}{2n}} \Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \sqrt{u/n}$$

To get the marginal pdf $f_T(t)$ the usual approach would be to compute

$$f_T(t) = \int_0^\infty f_{T,U}(t,u) \, du$$

$$= \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{n}} \frac{1}{\Gamma\left(\frac{n}{2}\right)} \int_0^\infty \frac{u^{n/2}-1}{\Gamma\left(\frac{n+1}{2}\right)} -e^{-\frac{t^2}{2n}} -\frac{u^2}{2n} \, du$$

The integral above can be re-written in terms of the gamma function. To see

$$I = \int_0^\infty u^{n/2-1} e^{-\frac{u^{2/n}+1}{2}} \, du \quad \text{c.o.v. let} \quad \lambda = u^{2/n} + 1$$

$$d\lambda = du \frac{t^{2/n}+1}{2}$$

$$= \int_0^\infty \left[2 \frac{e^{-\lambda}}{\lambda^{n/2}} \right]^{n/2-1} \lambda^{-\frac{n+1}{2}} e^{-\frac{2\lambda}{t^{2/n}+1}} \, d\lambda$$

$$= \left[\frac{2}{t^{2/n}+1}\right]^{n/2} \int_0^B \left[\frac{\lambda^{n+1}-1}{\lambda} e^{-\frac{\lambda}{2}} \right]^{n/2-1} e^{-\frac{2\lambda}{t^{2/n}+1}} \, d\lambda$$

$$= \left[\frac{2}{t^{2/n}+1}\right]^{n/2} \int_0^B \left[\frac{\lambda^{n+1}-1}{\lambda} e^{-\frac{\lambda}{2}} \right]^{n/2-1} e^{-\frac{2\lambda}{t^{2/n}+1}} \, d\lambda$$
Substituting the integral:

\[ f_T(t) = \frac{1}{\sqrt{2\pi} \sqrt{n}} \frac{1}{\Gamma(n/2)} \left[ \frac{2}{\frac{t^2}{n} + 1} \right]^{n+1/2} = \frac{\Gamma(n+1/2)}{\Gamma(n/2) \sqrt{\pi n}} \frac{1}{(1 + \frac{t^2}{n})^{n+1/2}} \quad t \in \mathbb{R}. \]

The F Distribution \( Y \sim \chi^2_p, Z \sim \chi^2_{n-p} \) and \( Y \perp Z \). Then

\[ F = \frac{Y/p}{Z/(n-p)} \]

is called an F statistic, aka an F random variable, or said to have an F distribution. The pdf is

\[ f_F(f) = \frac{\Gamma(n/2) [p/(n-p)]^{p/2}}{\Gamma(p/2) \Gamma((n-p)/2)} \left[ 1 + \left( \frac{p}{n-p} \right) f \right]^{-(n+p)/2} f > 0 \]

Derivation of F Distribution [Scharf p.176].

Define a transformation \((f,g) = W(y,z)\) by

\[ f = \frac{y/p}{z/(n-p)} \quad y, z > 0 \]
\[ g = z \]
The transformation is invertible and the inverse \( \mathbf{y}, \mathbf{z} = W^{-1}(f, g) \) is defined by
\[
\begin{align*}
  y &= f \frac{p}{n-p} \quad f, g > 0 \\
  z &= g
\end{align*}
\]

The Jacobian of the \( W^{-1} \) transformation is
\[
|J| = \det \begin{bmatrix}
  \frac{\partial y}{\partial f} & \frac{\partial y}{\partial g} \\
  \frac{\partial z}{\partial f} & \frac{\partial z}{\partial g}
\end{bmatrix} = \det \begin{bmatrix}
  g \frac{p}{n-p} & f \frac{p}{n-p} \\
  0 & 1
\end{bmatrix} = g \frac{p}{n-p}
\]

Then from the standard C.O.V. / transformation of rv formula,
\[
\mathcal{F}_{FG}(f, g) = \mathcal{F}_{YZ} \left( f \frac{p}{n-p}, g \right) \cdot g \frac{p}{n-p}
\]

By hypothesis this factors as the product of two chi-squared distributions i.e.
\[
\mathcal{F}_{YZ}(y, z) = \mathcal{F}_Y(y) \mathcal{F}_Z(z) \]
\[
= e^{-y/2} \frac{y^{(p/2)-1}}{2^{p/2} \Gamma(p/2)} \cdot \frac{e^{-z/2} z^{(n-p)/2-1}}{2^{(n-p)/2} \Gamma((n-p)/2)}
\]
Then to calculate the marginal distribution we integrate over the variable $g$:

$$
\phi(x) = \int \frac{\exp\left[-\frac{p}{n-p} \frac{p}{\frac{n}{2}} \left( \frac{q}{n-p} - 1 \right) \right]}{2^{p/2} \Gamma(p/2) \Gamma(n-p/2)} \frac{g^{p/2-1}}{\left(1 + \frac{p}{n-p} f \right)^{n/2}} \, dg
$$

$$
= \left(\frac{p}{n-p}\right)^{p/2-1} \int_{0}^{\infty} \frac{\exp\left[-\frac{p}{n-p} \frac{p}{\frac{n}{2}} \left( 1 + \frac{p}{n-p} f \right) \right]}{2^{p/2} \Gamma(p/2) \Gamma(n-p/2)} \frac{g^{p/2-1}}{\left(1 + \frac{p}{n-p} f \right)^{n/2}} \, dg
$$

\[ \text{\because Need to evaluate the integral} \]

$$
I = \int_{0}^{\infty} \exp\left\{-\frac{p}{n-p} \frac{p}{\frac{n}{2}} \left[1 + \frac{p}{n-p} f \right] g \right\} g^{n/2-1} \, dg
$$

for which we make a c.o.v. $w = \frac{p}{n-p} \left[1 + \frac{p}{n-p} f \right] g$

$\Rightarrow \quad dw = \frac{p}{n-p} \left[1 + \frac{p}{n-p} f \right] \, dg$

$\Rightarrow \quad I = \int_{0}^{\infty} e^{-w} w^{n/2-1} \left[\frac{p}{n-p} \left(1 + \frac{p}{n-p} f \right) \right]^{\frac{n}{2}} \, dw \left[\frac{p}{n-p} \left(1 + \frac{p}{n-p} f \right) \right]^{-1}$
\[
\Rightarrow I = \left[ \frac{1}{z} \left( 1 + \frac{p}{n-p} \right) f \right]^{-\frac{n}{2}} \int_0^\infty e^{-w} w^{\frac{n-p}{2}-1} \, dw
\]

\[
= \left[ \frac{1}{z} \left( 1 + \frac{p}{n-p} \right) f \right]^{-\frac{n}{2}} \Gamma\left( \frac{n}{2} \right)
\]

Substituting back in:

\[
\frac{\left( \frac{p}{n-p} \right) \left[ f \frac{p}{n-p} \right]^{\frac{n-p}{2}-1}}{2^{\frac{n}{2}} \Gamma\left( \frac{n}{2} \right) 2^{\frac{n-p}{2}} \Gamma\left( \frac{n-p}{2} \right)} \left[ \frac{1}{z} \left( 1 + \frac{p}{n-p} \right) f \right]^{-\frac{n}{2}} \Gamma\left( \frac{n}{2} \right)
\]

\[
= \frac{\Gamma\left( \frac{n}{2} \right) \left( \frac{p}{n-p} \right)^{\frac{n-p}{2}} 2^{\frac{n}{2}} \Gamma\left( \frac{n-p}{2} \right)}{\Gamma\left( \frac{n}{2} \right) \Gamma\left( \frac{n-p}{2} \right) 2^{\frac{n-p}{2}} 2^{\frac{n}{2}}} \frac{1}{{\left( 1 + \frac{p}{n-p} f \right)}^{\frac{n}{2}}}
\]

There are also non-central versions of \( \chi^2, T \) and \( F \) which might be useful at some point. See [Scharf pp. 176-178] or more general material on quadratic forms of Gaussian rvs.
Sample mean, sample standard deviation, and related properties. Let $X_1, X_2, \ldots$ be a sequence of real-valued random variables. The sample mean is defined to be

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i.$$ 

The sample variance is defined to be

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

(a) Define deviations from the sample mean $\tilde{X}_i \equiv X_i - \bar{X}_n, i = 1, 2, \ldots, n$. Show that only $n - 1$ of the deviations can be picked independently. For this reason one says that $S_n^2$ has $n - 1$ degrees of freedom.

(b) Assume that the $X_1, X_2, \ldots$ are i.i.d. with mean $\mu$ and variance $\sigma^2$. Find the mean and variance of the sample mean $\bar{X}_n$.

(c) Find the mean of $S_n^2$.

(d) Assume that the $X_1, X_2, \ldots$ are i.i.d. distributed as $\mathcal{N}(\mu, \sigma^2)$. Here the goal is to characterize the joint distribution of $\bar{X}_n$ and $S_n^2$.

i. Show that $\bar{X}_n \sim \mathcal{N}(\mu, \sigma^2/n)$.

ii. Show that $\bar{X}_n$ is statistically independent of the deviations $\tilde{X}_i$, for $i = 1, 2, \ldots, n$.

Hint: First establish the lemma:

Lemma: $X_1, X_2, \ldots, X_n$ i.i.d. of variance $\sigma^2$. Then

$$\operatorname{Cov}(\tilde{X}_i, \bar{X}_n) = 0$$

for $i = 1, 2, \ldots, n$.

and use the fact that uncorrelated $\Rightarrow$ statistically independent for Gaussians.

iii. Use the previous to show that $\bar{X}_n$ is statistically independent of $S_n^2$.

iv. Finally show that

$$\frac{(n-1)S_n^2}{\sigma^2} \sim \chi^2_{n-1}$$

which completely characterizes the joint distribution of $\bar{X}_n$ and $S_n^2$. 
Sample Mean, Standard Deviation, and Related Properties

\[ X_1, X_2, \ldots \]

a sequence of real-valued random variables. The sample mean is defined to be

\[ \overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \]

or

\[ \overline{X}_n \]

Clearly it is itself a random variable. The sample variance is defined to be

\[ S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2 \]

or

\[ S_n^2 \]

We defined the above without assumptions on the distribution of the \( X_i \). One important observation about the degrees of freedom is true for any distributions. Define deviations from the sample mean

\[ \tilde{X}_i \triangleq X_i - \overline{X} \]

\[ \Rightarrow \sum_{i=1}^{n} \tilde{X}_i = 0 \]

is the deviations cannot be picked independently. ... or only \( n-1 \) can be.
Re: Only \( n-1 \) degrees of freedom in \( n \) deviations from the sample mean.

\[
\sum_{i=1}^{n} X_i = \sum_{i=1}^{n} (X_i - \overline{X}) = \sum_{i=1}^{n} X_i - n \overline{X} = 0
\]

Properties of Sample Mean/Variance Assuming the \( X_i \) are iid with mean \( \mu \) and variance \( \sigma^2 \)

- \( \mathbb{E} \overline{X} = \mu \)
- \( \text{Var} \overline{X} = \text{Var} \left\{ \frac{1}{n} \sum_{i=1}^{n} X_i \right\} = \frac{1}{n^2} \text{Var} \left\{ \sum_{i=1}^{n} X_i \right\} = \frac{1}{n^2} \text{Cov} \left( \sum_{i=1}^{n} X_i, \sum_{j=1}^{n} X_j \right) \)
  \[
  = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \text{Cov} (X_i, X_j) \\
  = \frac{1}{n^2} \sum_{i=1}^{n} \text{Var} (X_i) + \frac{1}{n^2} \sum_{i \neq j} \text{Cov} (X_i, X_j) \\
  = \frac{n \sigma^2}{n^2} + \frac{1}{n^2} \sum_{i \neq j} \text{Cov} (X_i, X_j) \\
  = \frac{\sigma^2}{n} + 0 \text{ since indep} \]

- \( \text{Var} \overline{X} = \frac{\sigma^2}{n} \) by identical distributions
\[ (n-1) S^2 = \sum_{i=1}^{n} (X_i - \bar{X})^2 = \sum_{i=1}^{n} (X_i - \mu - \bar{X})^2 \]

\[ = \sum_{i=1}^{n} (X_i - \mu)^2 + \sum_{i=1}^{n} (\bar{X} - \mu)^2 - 2(\bar{X} - \mu) \sum_{i=1}^{n} (X_i - \mu) \]

\[ = \sum_{i=1}^{n} (X_i - \mu)^2 + n(\bar{X} - \mu)^2 - 2n(\bar{X} - \mu) \cdot n(\bar{X} - \mu) \]

\[ = \sum_{i=1}^{n} (X_i - \mu)^2 - n(\bar{X} - \mu)^2 \]

Take expectations of both sides to conclude

\[ (n-1) E(S^2) = n\sigma^2 - n \text{Var}(\bar{X}) \]

\[ = n\sigma^2 - n\sigma^2/n = (n-1)\sigma^2 \]

\[ \Rightarrow E(S^2) = \sigma^2 \]
The Joint Distribution of $\bar{X}$ and $S^2$ when $X_i, i=1,2,...,n$ is iid and $\sim N(\mu, \sigma^2)$

Obvious from prev. and fact that linear combinations of Normal rvs are Normal that

$$\bar{X} \sim N(\mu, \sigma^2/n)$$

Lemma: $X_1, \ldots, X_n$ iid of variance $\sigma^2$. Then

$$\text{Cov} \left( X_i - \bar{X}, \bar{X} \right) = 0 \quad i = 1, 2, \ldots, n$$

Compute

$$\text{Cov} \left( X_i - \bar{X}, \bar{X} \right) = \text{Cov} \left( X_i, \bar{X} \right) - \text{Cov} \left( \bar{X}, \bar{X} \right)$$

$$= \text{Cov} \left( X_i, \frac{1}{n} \sum_{j=1}^{n} X_j \right) - \text{Var} \left( \bar{X} \right)$$

$$= \frac{1}{n} \sum_{j=1}^{n} \text{Cov} \left( X_i, X_j \right) - \frac{\sigma^2}{n}$$

$$= \left\{ \begin{array}{ll} \sigma^2 & \text{for } j = i \\ 0 & \text{else} \end{array} \right\} - \frac{\sigma^2}{n} = 0.$$

\[\square\]

The random variables $\bar{X}, X_i, i=1,2,...,n$ are clearly jointly Gaussian from well-known properties. Then from the lemma and the fact that uncorrelated $\implies$ independent for Gaussians

$$\bar{X} \perp X_i - \bar{X} \quad i = 1, 2, \ldots, n$$

$$\implies \quad \bar{X} \perp \text{any function} \left( X_1 - \bar{X}, X_2 - \bar{X}, \ldots, X_n - \bar{X} \right)$$
\[ \bar{X} \perp S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2 \]

Thus have marginal dist of \( \bar{X} \), Stat. indp. of \( \bar{X} \) and \( S^2 \) \( \Rightarrow \) joint pdf is product of the two marginal pdfs

\[ (n-1)S^2 = \sum_{i=1}^{n} (X_i - \bar{X})^2 = \sum_{i=1}^{n} (X_i - \mu)^2 - n(\bar{X} - \mu)^2 \]

as in proof on p. c

\[ \Rightarrow \frac{(n-1)S^2}{\sigma^2} = \sum_{i=1}^{n} \left( \frac{X_i - \mu}{\sigma} \right)^2 - n \left( \frac{\bar{X} - \mu}{\sigma} \right)^2 \]

\[ \Rightarrow \frac{(n-1)S^2}{\sigma^2} + \left( \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \right)^2 = \sum_{i=1}^{n} \left( \frac{X_i - \mu}{\sigma} \right)^2 \]

\( \chi^2 \) with 1 degree of freedom

Sum of squares of \( n \) iid \( \mathcal{N}(0,1) \) rvs.

\( \Rightarrow \chi^2 \) with \( n \)-degrees of freedom
Characteristic Functions Can Help

Characteristic Function of a rv defined \( \Phi_x(w) = \mathbb{E}\{e^{iwX}\} \).
(The Fourier Transform of the pdf).

Calculating the char. funct. for \( X^2 \) is easy. First, assume that \( Y_1, \ldots, Y_n \) are iid \( \sim N(0,1) \). Define

\[ Z = Y_1^2 + Y_2^2 + \cdots + Y_n^2 \]

\[ \Phi_Z(w) = \mathbb{E}\left\{e^{iwZ}\right\} = \mathbb{E}\left\{\prod_{i=1}^n e^{iwy_i^2}\right\} = \left[\mathbb{E}\{e^{iwy_i^2}\}\right]^n \]

Since indep and identically distributed.

\[ \Phi_{Y_1}(w) = \mathbb{E}e^{iwY_1^2} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-y_1^2/2} e^{iwY_1^2} dy_1 \]

\[ = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-y_1^2(1/2 - jw)} dy_1 \]

Write \( 1/2 - j\omega = 1/2 e^{-\omega} \Rightarrow e^2 = (1/2 - j\omega)^{-1} \). Then from form of Gaussian pdf

\[ \frac{1}{\sqrt{\sigma^2}} \text{ Integral } = 1 \]

\[ \Rightarrow \text{ Integral } = \sigma = \frac{1}{\sqrt{2}} \left(\frac{1}{2} - j\omega\right)^{-\frac{1}{2}} \]

\[ = (1 - j2\omega)^{-\frac{1}{2}} \]
\[ \Phi_z(\omega) = \left(1 - j2\omega\right)^{-n/2} \]

Re: The Distribution of \( \frac{(n-1)S^2}{\sigma^2} \)

Since \( \frac{(n-1)S^2}{\sigma^2} \) \( \sim \) \( \left(\frac{X - \mu}{\sigma / \sqrt{m}}\right)^2 \) and since the char. funct. of a sum of indep. rvs is product of indiv. char. functs:

\[
\left( \text{char. funct of } \frac{(n-1)S^2}{\sigma^2} \right) \cdot (1 - j2\omega)^{-1/2} = (1 - j2\omega)^{-n/2}
\]

\[ \Rightarrow \text{char. funct. of } \frac{(n-1)S^2}{\sigma^2} = \left(1 - j2\omega\right)^{-(n-1)/2} \]

\[ \therefore \text{It is } \chi^2 \text{ of } n-1 \text{ degrees of freedom.} \]