Trabalho IV: Testes uniformemente mais poderosos.

Disciplina: Inferência Estatística
Aluno: Rener de Souza Oliveira

18 de Novembro de 2020

Introdução

Vimos que os testes de hipótese fornecem uma abordagem matematicamente sólida para traduzir hipóteses científicas sobre o processo gerador dos dados em decisões sobre os dados – isto é, traduzir afirmações sobre partições do espaço de parâmetros, Ω, em afirmações testáveis sobre o espaço amostral $\mathcal{X}^n$.
Um teste $\delta(\boldsymbol{X})$ é uma decisão (binária) de rejeitar ou não uma hipótese nula (H_0) sobre $\theta \in \Omega$ com base em uma amostra $\boldsymbol{X}$. A capacidade de um teste de rejeitar H_0 quando ela é falsa é medida pela função poder, $\pi(\theta|\delta)$. Nem todos os testes, no entanto, são criados iguais. Em certas situações, é possível mostrar que um procedimento δ_A é *uniformemente* mais poderoso que outro procedimento δ_B para testar a mesma hipótese.
Neste trabalho, vamos definir e aplicar o conceito de **teste uniformemente mais poderoso**.

1 - Motivação e Definição

Sejam:

$$\begin{aligned} H_0 :&\theta \in \Omega_0 \subset \Omega, \\ H_1 :&\theta \in \Omega_1 \subset \Omega, \\ \text{onde } &\Omega_1 = \Omega \setminus \Omega_0 \end{aligned} \tag{1}$$

Ao realizar um procedimento de teste $\delta(\boldsymbol{X})$, é desejável que a função poder $\pi(\theta|\delta) :\stackrel{\text{def}}{=} \Pr(\textit{Rejeitar } H_0|\theta)$ seja menor ou igual à um nível de significância $\alpha_0 \in (0,1)$, quando $\theta \in \Omega_0$, limitando superiormente a probabilidade de erro do tipo I (rejeitar H_0 quando ela é verdadeira). Podemos expressar tal propriedade da seguinte forma:

$$\alpha(\delta) \leq \alpha_0$$

Onde $\alpha(\delta) :\stackrel{\text{def}}{=} \sup_{\theta \in \Omega_0} \pi(\theta|\delta)$ é o tamanho do teste.

Além disso, queremos também ter algum controle sobre a probabilidade de erro do tipo II (não rejeitar H_0 quando ela é falsa). Como a probabilidade de tal erro quando $\theta \in \Omega_1$ é igual a $1 - \pi(\theta|\delta)$, queremos que, na região onde H_0 é falsa (Ω_1) a função poder $\pi(\theta|\delta)$ seja máxima, para todo θ em tal região. Tal maximização, minimiza a probabilidade de erro do tipo II quando $\theta \in \Omega_1$, isso nem sempre é possível, mas quando for, temos um nome especial para esse teste, que segue abaixo sua definição:

Definicão 1 ***(Teste Uniformemente mais poderoso)*** *Seja $\mathcal{C}$ uma classe de teste para as hipóteses (1); $\delta^* \in \mathcal{C}$ é chamado de uniformemente mais poderoso (UMP[1]) da classe $\mathcal{C}$, se:*

$$\pi(\theta|\delta^*) \geq \pi(\theta|\delta) \; \forall \, \theta \in \Omega_1,$$

para qualquer teste $\delta \in \mathcal{C}$.

Seguindo a motivação dada acima, podemos definir $\mathcal{C}$ como o conjunto de todos dos testes de tamanho menor ou igual a α_0, limitando o erro tipo I. Neste caso, chamamos δ^* de UMP para (1) ao nível α_0.

2 - Razão de Verossimilhança Monótona

Definicão 2 ***(Razão de Verossimilhanças Monótona)*** *Seja $f_n(\boldsymbol{x}|\theta)$ a função de verossimilhança das observações $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$, e $T = r(\boldsymbol{X})$ uma estatística. Dizemos que a distribuição dos dados tem* ***razão de verossimilhanças monótona*** *sob T, quando, $\forall \, \theta_1, \theta_2 \in \Omega; \theta_1 < \theta_2$, a razão $\dfrac{f(\boldsymbol{x}|\theta_2)}{f(\boldsymbol{x}|\theta_1)}$ depende dos dados através de $r(\boldsymbol{x})$ somente, e é uma função monótona de $r(\boldsymbol{x})$ sob seu espaço de definição.*

3 - UMP para H_0 simples

Considere uma hipótese nula simples, $H_0 : \theta = \theta_0$, $\theta_0 \in \Omega$. Mostraremos que, se vale o Teorema da Fatorização, e existem c e α_0 tais que

$$\Pr\left(r(\boldsymbol{X}) \geq c \mid \theta = \theta_0\right) = \alpha_0,$$

então o procedimento δ^* que rejeita H_0 se $r(\boldsymbol{X}) \geq c$ é UMP para H_0 ao nível α_0.
Mas antes, vamos enunciar alguns teoremas:

Teorema 1 ***(Teorema da Fatorização)[3]*** *Sejam $X_1, X_2, \ldots, X_n$ amostra aleatória de uma distribuição de densidade ou massa $f(x|\theta)$, onde $\theta \in \Omega$. Uma estatística $T = r(X_1, X_2, \ldots, X_n)$ é suficiente para θ, se, e somente se a distribuição conjunta dos dados $f_n(\boldsymbol{x}|\theta)$ pode ser fatorizada como:*

$$f_n(\boldsymbol{x}|\theta) = u(\boldsymbol{x}) v[r(\boldsymbol{x}), \theta],$$

para todo $\boldsymbol{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, e $\forall \theta \in \Omega$. u e v são funções não negativas.

[1] Uniformly Most Powerful Test

A demostração pode ser encontrada em [3].

Teorema 2 ***(Lema de Neyman-Pearson)[1]*** *Seja* $(X_1, X_2, \ldots, X_n) \in \mathbb{R}^n$ *uma amostra indexada por* θ*. Considere as hipóteses*

$$\begin{aligned} H_0 &: \theta = \theta_0, \\ H_1 &: \theta = \theta_1, \end{aligned} \tag{2}$$

e seja $f_n(\boldsymbol{x}|\theta_i)$*, com* $i = 0, 1$ *a função de densidade ou massa dos dados.*
Seja $R \in \mathbb{R}^n$ *uma região de rejeição que satisfaça:*

$$\begin{gathered} \boldsymbol{x} \in R \text{ se } f(\boldsymbol{x}|\theta_1) \geq kf(\boldsymbol{x}|\theta_0) \\ \text{e } \boldsymbol{x} \in R^C \text{ se } f(\boldsymbol{x}|\theta_1) \leq kf(\boldsymbol{x}|\theta_0), \end{gathered} \tag{3}$$

para algum $k \geq 0$ *e*

$$\Pr(\boldsymbol{X} \in R|\theta = \theta_0) = \alpha_0. \tag{4}$$

Então, todo teste que satisfaz (3) e (4) é UMP ao nivel α_0*.*

A demostração será omitida pois pode ser encontrada em [1].

Corolário 1 *Considere as hipóteses (2). Seja* $T(\boldsymbol{X})$ *uma estatística suficiente para* θ *e* $g(t|\theta_i)\, i = 0, 1$*, uma função de* $t = T(\boldsymbol{x})$ *tal que fatoriza a verossimilhança dos dados em* $f_n(\boldsymbol{x}|\theta_i) = g(t|\theta_i)u(\boldsymbol{x})$*, para alguma função* $u(\boldsymbol{x}) \geq 0$*.*
Seja δ *um teste que rejeite* H_0 *se* T *pertence a uma região de rejeição* S *(subconjunto do espaço de definição de* T*). Assim,* δ *será UMP ao nível* α_0 *se satisfazer:*

$$\begin{gathered} g(t|\theta_1) \geq kg(t|\theta_0) \implies t \in S \\ \text{e } g(t|\theta_1) \leq kg(t|\theta_0) \implies t \in S^C, \end{gathered} \tag{5}$$

para algum $k \geq 0$ *e*

$$\Pr[T(\boldsymbol{X}) \in S|\theta = \theta_0] = \alpha_0. \tag{6}$$

Demonstração: Definindo $R = \{\boldsymbol{x}|T(\boldsymbol{x}) \in S\}$, rejeitaremos H_0 se $\boldsymbol{x} \in R$. Pelo Teorema da Fatorização, dado que $T(\boldsymbol{X})$ é suficiente, a verossimilhança de $\boldsymbol{X}$ pode ser escrita como $f_n(\boldsymbol{x}|\theta_i) = g(T(\boldsymbol{x})|\theta_i)u(\boldsymbol{x})$, $i = 0, 1$, para alguma função $u(\boldsymbol{x}) \geq 0$.
Multiplicando tal função nas desigualdades (5) temos:

$$\begin{aligned} & g(T(\boldsymbol{x})|\theta_1) \geq kg(T(\boldsymbol{x})|\theta_0) \\ \Leftrightarrow & g(T(\boldsymbol{x})|\theta_1)u(\boldsymbol{x}) \geq kg(T(\boldsymbol{x})|\theta_0)u(\boldsymbol{x}) \\ \Leftrightarrow & f_n(\boldsymbol{x}|\theta_1) \geq kf_n(\boldsymbol{x}|\theta_0) \end{aligned}$$

Assim, tem-se: $f_n(\boldsymbol{x}|\theta_1) \geq kf_n(\boldsymbol{x}|\theta_0) \implies T(\boldsymbol{x}) \in S \implies \boldsymbol{x} \in R$.
Analogamente, $f_n(\boldsymbol{x}|\theta_1) \leq kf_n(\boldsymbol{x}|\theta_0) \implies \boldsymbol{x} \in R^C$.
De (6), tem-se:

$$\Pr(\boldsymbol{X} \in R|\theta = \theta_0) = \Pr[T(\boldsymbol{X}) \in S|\theta = \theta_0] = \alpha_0$$

Pelo **Lema de Neyman-Pearson** concluímos que o teste δ é UMP ao nível α_0.
Voltando agora ao problema inicial da seção, queremos provar que δ^* é UMP ao nível α_0 para $H_0 : \theta = \theta_0$.
Primeiramente precisamos provar que $\alpha(\delta^*) = \alpha_0$.

$$\begin{aligned}\alpha(\delta^*) &= \sup_{\theta \in \Omega_0} \pi(\theta|\delta^*)\\ &= \sup_{\theta \in \Omega_0} \Pr[r(\boldsymbol{X}) \geq c|\theta]\end{aligned}$$

Como $\Omega_0 = \{\theta_0\}$, o supremo ocorre em θ_0 o que implica que $\alpha(\delta^*) = \alpha_0$.
Agora precisamos provar que δ^* é UMP.
Façamos θ' arbitrário, com $\theta' \neq \theta_0$, testaremos $H_0 : \theta = \theta_0$ contra $H_1' : \theta = \theta'$.
No problema em questão, vale o **Teorema da Fatorização** para $r(\boldsymbol{X})$, logo assumindo sua suficiência, temos que a verossimilhança pode ser escrita como $f_n(\boldsymbol{x}|\theta) = g(r(\boldsymbol{x})|\theta)u(\boldsymbol{x})$, para alguma função $u(\boldsymbol{x}) \geq 0$.
Seja $t = r(\boldsymbol{x})$; Definamos:

$$k = \inf_{t \in \mathcal{T}} \frac{f_n(\boldsymbol{x}|\theta')}{f_n(\boldsymbol{x}|\theta_0)} = \frac{g(t|\theta')}{g(t|\theta_0)}$$

Com $\mathcal{T} :\overset{\text{def}}{=} \{t|t \geq c\}$
Tal ínfimo existe, pois pelo Teorema da Fatorização, a função g é não-negativa, logo, o conjunto na qual estamos tomando ínfimo é limitado inferiormente por 0. Pelo análogo do Axioma do Supremo para ínfimos, k está bem definido.
Pela definição de ínfimo segue que:
$r(\boldsymbol{x}) \geq c \Leftrightarrow \dfrac{g(r(\boldsymbol{x})|\theta')}{g(r(\boldsymbol{x})|\theta_0)} \geq k$.
Pelo Corolário 1 do Lema de Neyman-Pearson, temos que δ^* é UMP para as hipóteses $H_0 : \theta = \theta_0$ e $H_1' : \theta = \theta'$, ou seja, $\pi(\theta|\delta^*) \geq \pi(\theta'|\delta)$, para qualquer teste δ de tamanho α_0. Como θ' foi escolhido arbitrariamente diferente de θ_0, temos que δ^* satisfaz $\pi(\theta|\delta^*) \geq \pi(\theta'|\delta)\ \forall \theta' \neq \theta_0$, o que prova nossa afirmação inicial. ■

4 - Duas-Caras e UMP para Bernoulli

Suponha que você encontra o Duas-Caras na rua e ele não vai com a sua... cara. Ele decide jogar a sua famosa moeda para o alto para decidir se te dá um cascudo. Se der cara (C), você toma um cascudo. Você, que sabe bem Estatística, pede que ele pelo menos jogue a moeda umas $n = 10$ vezes antes de tomar a decisão derradeira.

Surpreendentemente, ele concorda. Lança a moeda e obtém

KCKCKCCKKK

Você agora deve decidir se foge, se arriscando a tomar dois cascudos ao invés de um, ou se fica e possivelmente não toma cascudo nenhum. Se p é a probabilidade de dar cara, estamos interessados em testar a hipótese

$$H_0 : p \leq \frac{1}{2},$$
$$H_1 : p > \frac{1}{2}.$$

1. Escreva a razão de verossimilhanças para esta situação;

   Sejam p_0 e p_1, tais que $0 < p_0 \leq \frac{1}{2} < p_1 < 1$. Seja X_i a variável indicadora de cara no i-ésimo lançamento do Duas-Caras; Assumindo que os 10 lançamentos são independentes, temos $X_1, X_2, \ldots, X_{10} \overset{iid}{\sim}$ Bernoulli(p), na qual $f(x_i|p) = p^{x_i}(1-p)^{1-x_i}$.

   A verossimilhança sera então:

   $$f_n(\boldsymbol{x}|p) = \prod_{i=1}^{10} p^{x_i}(1-p)^{1-x_i} = p^y(1-p)^{10-y},$$

   onde $y = \sum_{i=1}^{10} x_i$.

   Assim, a razão de verossimilhança será:

   $$\begin{aligned}\frac{f_n(\boldsymbol{x}|p_1)}{f_n(\boldsymbol{x}|p_0)} &= \frac{p_1^y(1-p_1)^{10-y}}{p_0^y(1-p_0)^{10-y}} \\ &= \left[\frac{p_1(1-p_0)}{p_0(1-p_1)}\right]^y \left(\frac{1-p_1}{1-p_0}\right)^{10} \qquad (7)\end{aligned}$$

   Podemos ver que a razão depende dos dados somente através da estatística suficiente y, e que a expressão é monótona em y, pois $p_0 < p_1 \Rightarrow \frac{p_1(1-p_0)}{p_0(1-p_1)} > 1$, que mostra que a razão é estritamente crescente neste caso. Por definição, dizemos que a distribuição dos dados tem razão de verossimilhança monótona crescente sob y (MLR[2] crescente).

2. Nesta situação, é do seu interesse encontrar um teste UMP. Faça isso e aplique o teste desenvolvido aos dados que conseguiu arrancar do Duas-Caras.

   Existe uma generalização dos resultados da seção anterior, que estende a noção de existência de UMP para H_0 composta, e inclui a hipótese de distribuição com MLR. Vamos enunciá-lo e demonstrá-lo brevemente, pois basta algumas adaptações da demonstração da seção anterior para H_0 simples.

[2] Monotone Likelihood Ratio

Teorema 3 *(Teorema de Karlin-Rubin)[2] Sejam as hipóteses:*

$$\begin{aligned} H_0 &: \theta \leq \theta_0 \\ H_1 &: \theta > \theta_0 \end{aligned} \tag{8}$$

Seja $T = r(\boldsymbol{X})$ uma estatística suficiente para θ e suponha que a família de distribuições dos dados $\{f(\boldsymbol{x}|\theta)|\theta \in \Omega\}$ tem razão de verossimilhança monótona não-decrescente sob T. Assim, para qualquer c, o teste δ^* que rejeita H_0 se $T \geq c$ é um teste UMP ao nível α_0, onde $\alpha_0 = \Pr[T \geq c\,|\theta = \theta_0]$.

Demonstração: Devemos mostrar que o tamanho do teste é α_0 e que δ^* é UMP. O fluxo da demostração será parecido com a do Corolário (1).

Para a primeira parte, queremos que,

$$\sup_{\theta \leq \theta_0} \Pr[T \geq c\,|\theta] = \alpha_0$$

Na última demostração o supremo era trivial pois era tomado em um conjunto unitário, agora a situação se complica. Se a função poder for não-decrescente, o supremo ocorrerá em θ_0 e fica provada essa parte, pois foi suposto que $\Pr[T \geq c\,|\theta = \theta_0] = \alpha_0$. Provaremos isso ao longo deste texto.

Fixemos θ' arbitrário tal que $\theta' > \theta_0$, testaremos $H_0' : \theta = \theta_0$ contra $H_1' : \theta = \theta'$. Como $T = r(\boldsymbol{X})$ é suficiente, pelo **Teorema da Fatorização**, temos que a verossimilhança pode ser escrita como $f_n(\boldsymbol{x}|\theta) = g(t|\theta)u(\boldsymbol{x})$, para alguma função $u(\boldsymbol{x}) \geq 0$, com $t = r(\boldsymbol{x})$.

Definamos:

$$k = \inf_{t \in \mathcal{T}} \frac{f_n(\boldsymbol{x}|\theta')}{f_n(\boldsymbol{x}|\theta_0)} = \frac{g(t|\theta')}{g(t|\theta_0)}$$

Com $\mathcal{T} \stackrel{\text{def}}{:=} \{t|t \geq c\}$. A existência de tal ínfimo já foi justificada.

Segue que:

$t \geq c \Leftrightarrow \dfrac{g(t|\theta')}{g(t|\theta_0)} \geq k.$

[5]Pelo Corolário (1), temos que δ^* é UMP para a hipótese simples H_0', ou seja, $\pi(\theta'|\delta^*) \leq \pi(\theta'|\delta_2)$, onde δ_2 é um teste qualquer de nível α_0 para H_0'. Como o processo não depende de θ', δ^* é UMP considerando $H_1 : \theta > \theta_0$.

O poder de δ^* avaliado em $\theta = \theta_0$ é igual a α_0 pelo enunciado do teorema. Para qualquer $\theta' > \theta_0$, temos $\pi(\theta'|\delta^*) \leq \alpha_0$, pelo fato de ser mais poderoso que um teste de tamanho igual a α_0 por exemplo.

Tomando um $\theta_2 < \theta_0$, e testando $H_0'' : \theta = \theta_2$ contra $H_1'' : \theta > \theta_2$. O teste δ^* terá um certo tamanho α''. por argumentos similares aos apresentados acima, tal teste é mais poderoso (UMP) que qualquer outro de nível α'', como o poder avaliado em $\theta = \theta_0$ é α_0, conclui-se que $\alpha'' \leq \alpha_0$. Dessa forma, vemos que a função poder é não-decrescente em θ. Assim concluímos o argumento inicial, de que $\sup_{\theta \leq \theta_0} \pi(\theta|\delta^*)$.

A conclusão de que δ^* é UMP para as hipóteses (8) segue do argumento de que, dado um teste δ de tamanho $\alpha(\delta) \leq \alpha_0$ e seu poder não excede $\pi(\theta|\delta^*)$, pois δ tem poder menor ou igual a α_0, quando avaliado em $\theta = \theta_0$, por continuidade e monotinicidade, o poder de δ será menor que o de δ^*, o que finaliza a prova. ■

Voltando ao problema do duas caras, temos todas as condições satisfeitas para aplicar **Karlin-Rubin**: As hipóteses são do mesmo formato, y é estatística suficiente, a razão de verossimilhança é crescente sob y. Assim escolheremos um limiar c de forma forçada para que $\Pr(y \geq c|p = 1/2) = \alpha_0$, ou seja $1 - F(c|p = 1/2) = \alpha_0 \Rightarrow F(c|p = 1/2) = 1 - \alpha_0$, onde $F(x|p)$ é a distribuição acumulada da Binomial$(10, p)$.

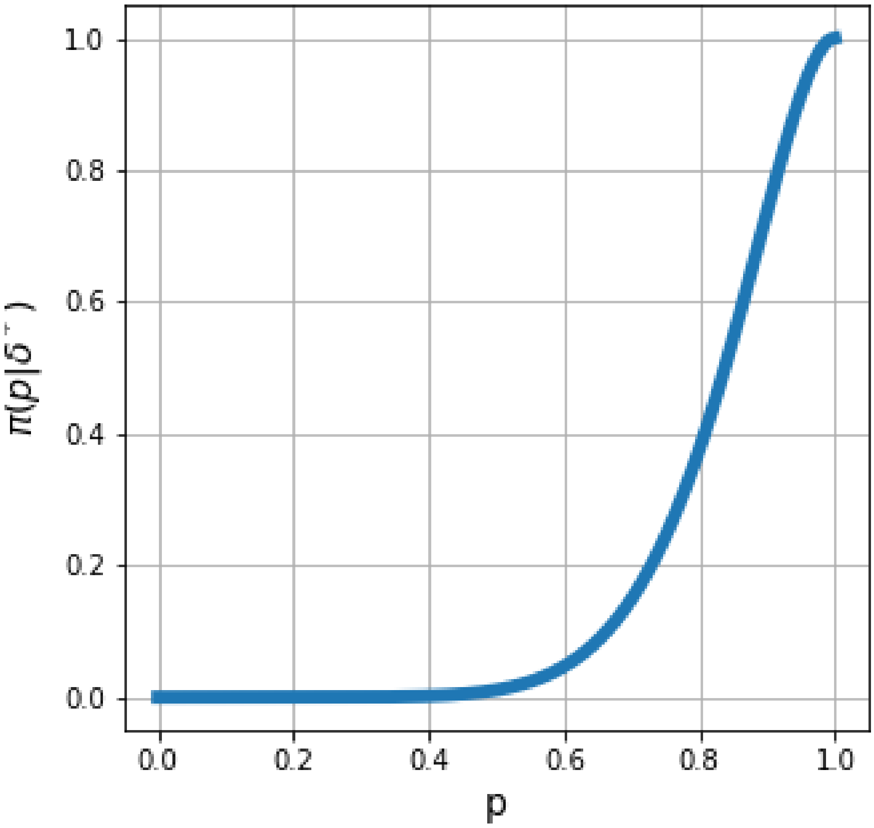
c	$F(c\|p = 1/2)$
0	0.000977
1	0.010742
2	0.054688
3	0.171875
4	0.376953
5	0.623047
6	0.828125
7	0.945312
8	0.989258
9	0.999023
10	1.000000

Tabela 1: Tabela gerada com Python 3.7

Vamos tomar $\alpha_0 = 0.05$, assim $1 - \alpha_0 = 0.95$. Como a distribuição é discreta, esses valores são são atingíveis nessa formulação, escolhemos então c tal que o poder é $\leq \alpha_0$, ou seja $F(c) \geq 0.95$. No caso faremos $c = 8$ pela Tabela 1.

No nosso caso o y observado foi $y = 5$ caras. Como $y < c$, falhamos em rejeitar a hipótese nula $p \leq 1/2$, o que empresta credibilidade para tal. Sendo assim, é razoável aceitar a proposta do Duas-Caras de lançar a moeda, pois no pior dos casos, tomaremos um cascudo apenas, ao invés de dois, caso fugíssemos.

Veja abaixo um gráfico da função poder quando $c = 8$ em função de p. Veja que realmente é uma função não-decrescente como havíamos demonstrado.



3. (Bônus) Mostre que, no item anterior, não é possível atingir qualquer nível α_0, isto é, que α_0 toma um número finito de valores. Proponha uma solução para que seja possível atingir qualquer nível em $(0, 1)$.

   [4] Como vimos pela tabela, o fato da distribuição binomial ser discreta, impede que seja possível atingir qualquer α_0. Uma solução alternativa, soluciona isso facilmente:

   No teste modificado, se $y < c$, rejeita-se H_0, se $y > c$ falha-se em rejeitar H_0, mas se $y = c$, rejeitamos H_0 com probabilidade q de tal forma que:

   $\Pr(y > c|p = 1/2) + q \Pr(y = c|p = 1/2) = \alpha_0$

   Assim, fazendo $q = \dfrac{\alpha_0 - \Pr(y > c|p = 1/2)}{\Pr(y = c|p = 1/2)}$ chegamos ao nível desejado.

   No caso acima, com c = 8, $\alpha_0 = 0.05$, temos $q = 0.89\overline{3}$ (Python 3.7).

Bibliografia

[1] George Casella and Roger Berger. *Statistical Inference*, pages 388–389. Duxbury Resource Center, June 2001.

[2] George Casella and Roger Berger. *Statistical Inference, Theorem 8.3.7*, pages 391–392. Duxbury Resource Center, June 2001.

[3] M.H. DeGroot and M.J. Schervish. *Probability and Statistics, 4th ed.*, pages 444–446. Addison-Wesley, 2012.

[4] M.H. DeGroot and M.J. Schervish. *Probability and Statistics, 4th ed.*, pages 556–557. Addison-Wesley, 2012.

[5] L.L. Scharf and C. Demeure. *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*, pages 124–126. Addison-Wesley series in electrical and computer engineering. Addison-Wesley Publishing Company, 1991.