STATISTICAL INFERENCE

- A Quick Guide -

Huan Q. Bui

Colby College

PHYSICS & MATHEMATICS
Statistics

Class of 2021

February 7, 2020

Preface

Greetings,

This guide is based on SC482: Statistical Inference, taught by Professor Liam O'Brien. The guide consists of lecture notes and material from *Introduction to Mathematical Statistics, 8th edition* by Hogg, McKean, and Craig. A majority of the text will be reading notes and solutions to selected problems.

As this is intended only to be a reference source, I might not be as meticulous with my explanations as I have been in some other guides.

Enjoy!

Contents

Part 1

Special Distributions

1.1 The Binomial and Related Distributions

If we let the random variable X equal the number of observed successes in n independent Bernoulli trials, each with success probability of p, then X follows the binomial distribution.s

A binomial pmf is given by

$$\boxed{p(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & x = 0, 1, 2, \dots \\ 0, \quad \text{else} \end{cases}} \tag{1.1}$$

Using the binomial expansion formula, we can easily check that

$$\boxed{\sum_x p(x) = 1} \tag{1.2}$$

The mgf of a binomial distribution is obtained by:

$$\boxed{M_{\text{bin}}(t) = E[e^{tx}] = \sum_x e^{tx} p(x) = \left[(1-p) + pe^t\right]^n \, \forall t \in \mathbb{R}} \tag{1.3}$$

With this, we can find the mean and variance for $p(x)$:

$$\boxed{\mu = M'(0) = n, \quad \sigma^2 = M''(0) = np(1-p)} \tag{1.4}$$

Theorem: Let $X_1, X_2, \dots, X_m$ be independent binomial random variables such that $X_i \sim \text{bin}(n_i, p), i = 1, 2, \dots, m$. Then

$$\boxed{Y = \sum_{i=1}^m X_i \sim \text{bin}\left(\sum_{i=1}^m n_i, p\right)} \tag{1.5}$$

Proof: We prove this via the mgf for Y. By independence, we have that

$$M_Y(t) = \prod_{i=1}^m (1 - p + pe^t)^{n_i} = (1 - p + pe^t)^{\sum_{i=1}^m n_i} \tag{1.6}$$

The mgf completely determines the distribution which Y follows, so we're done. □

1.1.1 Negative Binomial & Geometric Distribution

Consider a sequence of independent Bernoulli trials with constant probability p of success. The random variable Y which denotes the total number of failures in this sequence before the rth success follows the negative binomial distribution.

A negative binomial pmf is given by

$$\boxed{p_Y(t) = \begin{cases} \binom{(y+r-1)}{r-1} p^r (1-p)^y & y = 0, 1, 2, \ldots \\ 0, & \text{else} \end{cases}} \tag{1.7}$$

The mgf of this distribution is

$$\boxed{M(t) = p^r[1-(1-p)e^t]^{-r}} \tag{1.8}$$

When $r = 1$, Y follows the geometric distribution, whose pmf is given by

$$\boxed{p_Y(y) = p(1-p)^y, \quad y = 0, 1, 2, \ldots} \tag{1.9}$$

The mgf of this distribution is

$$\boxed{M(t) = p[1-(1-p)e^t]^{-1}} \tag{1.10}$$

1.2 Multinomial Distribution

We won't worry about this for now.

1.3 Hypergeometric Distribution

We won't worry about this for now.

1.4 The Poisson Distribution

The Poisson distribution gives the probability of observing x occurrences of some rare events characterized by rate $\lambda > 0$. The pmf is given by

$$\boxed{p(x) = \begin{cases} \frac{\lambda^x e^{-\lambda}}{x!}, & x = 0, 1, 2, \ldots \\ 0, & \text{else} \end{cases}} \tag{1.11}$$

We say a random parameter with the pmf of the form of $p(x)$ follows the Poisson distribution with parameter λ.

The mgf of a Poisson distribution is given by

$$\boxed{M(t) = e^{-\lambda(e^t-1)}} \tag{1.12}$$

From here, we can find the mean and variance:

$$\boxed{\mu = M'(0) = \lambda, \quad \sigma^2 = M''(0) = \lambda} \tag{1.13}$$

Theorem: If $X_1, \ldots, X_n$ are independent random variables, each $X_i \sim \text{Poi}(\lambda_i)$, then

$$\boxed{Y = \sum_{i=1}^{n} X_i \sim \text{Poi}\left(\sum_{i=1}^{n} \lambda_i\right)} \tag{1.14}$$

Proof: We once again prove this via the mgf of Y:

$$M_Y(t) = \prod_{i=1}^{n} e^{\lambda_i(e^t-1)} = e^{\sum_{i=1}^{n} \lambda_i(e^t-1)} \tag{1.15}$$

□

1.5 The Γ, χ^2, β distributions

The gamma function of $\alpha > 0$ is given by

$$\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} e^{-y}\, dy, \tag{1.16}$$

which gives $\Gamma(1) = 1$ and $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$.

1.5.1 The Γ and exponential distribution

A continuous random variable $X \sim \Gamma(\alpha, \beta)$ where $\alpha > 0$ and $\beta > 0$ whenever its pdf is

$$\boxed{f(x) = \begin{cases} \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta}, & 0 < x < \infty \\ 0, & \text{else} \end{cases}} \tag{1.17}$$

The mgf for X is obtained via the change of variable $y = x(1 - \beta t)/\beta$, where $t < 1/\beta$:

$$\boxed{M(t) = \int_0^\infty \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x(1-\beta t)/\beta}\, dx = \frac{1}{(1-\beta t)^\alpha}} \tag{1.18}$$

From here, we can find the mean and variance:

$$\boxed{\mu = M'(0) = \alpha\beta, \quad \sigma^2 = \alpha\beta^2} \tag{1.19}$$

The $\Gamma(1, \beta)$ distribution is a special case, and it is called the **exponential distribution** with parameter $1/\beta$.

Theorem: Let $X_1, \ldots, X_n$ be independent random variables, with $X_i \sim \Gamma(\alpha_i, \beta)$. Then

$$\boxed{Y = \sum_{i=1}^{n} X_i \sim \Gamma\left(\sum_{i=1}^{n} \alpha_i, \beta\right)} \tag{1.20}$$

Proof: Can you guess via which device we prove the statement above? □

1.5.2 The χ^2 distribution

The χ^2 distribution is a special case of the gamma distribution where $\alpha = r/2, r \in \mathbb{N}^*$ and $\beta = 2$. If a continuous r.v. $X \sim \chi^2(r)$ then its pdf is

$$\boxed{f(x) = \begin{cases} \frac{1}{\Gamma(r/2)2^{r/2}} x^{r/2-1} e^{-x/2}, & 0 < x < \infty \\ 0, & \text{else} \end{cases}} \tag{1.21}$$

Its mgf is

$$\boxed{M(t) = (1-2t)^{-r/2}, \quad t < \frac{1}{2}} \tag{1.22}$$

Theorem: Let $X \sim \chi^2(r)$ and $k > -r/2$ be given. Then $E[X^k]$ exists and is given by

$$\boxed{E[X^k] = \frac{2^k \Gamma(r/2+k)}{\Gamma(r/2)}} \tag{1.23}$$

Proof: is proof is purey computational and is left to the reader. □

From here, we note that all moments of the χ^2 distribution exist.

Theorem: Let $X_1, \ldots, X_n$ be r.v. with $X_i \sim \chi^2(r_i)$. Then

$$\boxed{Y = \sum_{i=1}^{n} X_i \sim \chi^2\left(\sum_{i=1}^{n} r_i\right)} \tag{1.24}$$

Proof: we once again find the mgf for Y. □

1.5.3 The β distribution

The β distribution differs from the other continuous ones we've discussed so far because its support are bounded intervals.

I will skip most of the details here, except mentioning that we can derive the beta distribution from the a pair of independent Γ random variables. Suppose $Y = X_1/(X_1 + X_2)$ where $X_i \sim \Gamma(\alpha, \beta)$ then the pdf of Y is that of the beta distribution:

$$\boxed{g(y) = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} y^{\alpha-1}(1-y)^{\beta-1}, & 0 < y < 1 \\ 0, & \text{else} \end{cases}} \tag{1.25}$$

The mean and variance of Y are

$$\boxed{\mu = \frac{\alpha}{\alpha+\beta}, \quad \sigma^2 = \frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}} \tag{1.26}$$

1.6 The Normal distribution

I have dedicated a large chunk in the QFT notes to evaluating Gaussian integrals, so I won't go into that here.

$X \sim \mathcal{N}(\mu, \sigma^2)$ whenever its pdf is

$$\boxed{f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right), \quad -\infty < x < \infty} \tag{1.27}$$

where μ and σ^2 are the mean and variance of X, respectively.

The mgf of X is can be obtained via the substitution $X = \sigma Z + \mu$:

$$\boxed{M(t) = \exp\left(\mu t + \frac{1}{2}\sigma^2 t^2\right)} \tag{1.28}$$

We note the following correspondence for $X = \sigma Z + \mu$:

$$X \sim \mathcal{N}(\mu, \sigma^2) \iff Z \sim \mathcal{N}(0, 1) \tag{1.29}$$

Theorem: $X \sim \mathcal{N}(\mu, \sigma^2) \implies V = (X-\mu)^2/\sigma^2 \sim \chi^2(1)$, i.e. a standardized, squared normal follows a chi-square distribution.

Proof: The proof isn't too hard. Let us write V as W^2 and so $W \sim \mathcal{N}(0,1)$. We consider the cdf $G(v)$ for V, with $v \geq 0$:

$$G(v) = P(W^2 \leq v) = P(-\sqrt{v} \leq W \leq \sqrt{v}) = 2\int_0^{\sqrt{v}} \frac{1}{\sqrt{2\pi}} e^{-w^2/2}\, dw \tag{1.30}$$

with $G(v) = 0$ whenever $v < 0$. From here, we can see that the pdf for v, under the change of notation $w \to \sqrt{y}$, is

$$g(v) = G'(v) = \frac{d}{dv}\left\{\int_0^v \frac{1}{\sqrt{2\pi}\sqrt{y}} e^{-y/2}\, dy\right\}, \quad 0 \geq v \tag{1.31}$$

or 0 otherwise. This means

$$\boxed{g(v) = \begin{cases} \frac{1}{\sqrt{\pi}\sqrt{2}} v^{1/2-1} e^{-v/2}, & 0 < v < \infty \\ 0, & \text{else} \end{cases}} \tag{1.32}$$

Using the fact that $\Gamma(1/2) = \sqrt{\pi}$ and by verifying that $g(v)$ integrates to unity we show $V \sim \chi^2(1)$. $\square$

Theorem: Let $X_1, \ldots, X_n$ be independent r.v. with $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$. Then for constants $a_1, \ldots, a_n$

$$\boxed{Y = \sum_{i=1}^{n} a_i X_i \sim \mathcal{N}\left(\sum_{i=1}^{n} a_i \mu_i, \sum_{i=1}^{n} a_i^2 \sigma_i^2\right)} \tag{1.33}$$

Proof: We once again prove this kind of theorems via the mgf for Y:

$$\begin{aligned} M(t) &= \prod_{i=1}^{n} \exp\left(t a_i \mu_i + \frac{1}{2} a_i^2 \sigma_i^2\right) \\ &= \exp\left\{t \sum_{i=1}^{n} a_i \mu_i + \frac{1}{2} t^2 \sum_{i=1}^{n} a_i^2 \sigma_i^2\right\} \end{aligned} \tag{1.34}$$

which is the mgf for the normal with the corresponding mean and variance above. □

Corollary: Let $X_1, \ldots, X_n \sim \mathcal{N}(\mu, \sigma^2)$. Then

$$\boxed{\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} \sim \mathcal{N}\left(\mu, \sigma^2/n\right)} \tag{1.35}$$

Proof: the proof is left to the reader.

1.6.1 Contaminated Normal

We won't worry about this for now.

1.7 The Multivariate Normal

I'll just jump straight to the n-dimensional generalization. Evaluations of high-dimensional Gaussian integrals and moments can also be found in the QFT notes.

We say an n-dimensional random vector $\mathbf{X}$ has a multivariate normal distribution if its mgf is

$$\boxed{M_{\mathbf{X}}(t) = \exp\left(\mathbf{t}^\top \boldsymbol{\mu} + \frac{1}{2}\mathbf{t}^\top \boldsymbol{\Sigma} \mathbf{t}\right)} \tag{1.36}$$

for all $\mathbf{t} \in \mathbb{R}^n$, where $\boldsymbol{\Sigma}$ is a symmetric, positive semi-definite matrix and $\boldsymbol{\mu} \in \mathbb{R}^n$. For short, we say $\mathbf{X} \sim \mathcal{N}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

Theorem: Suppose $\mathbf{X} \sim \mathcal{N}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ where $\boldsymbol{\Sigma}$ is positive definite. Then

$$\boxed{\mathbf{Y} = (\mathbf{X} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{X} - \boldsymbol{\mu}) \sim \chi^2(1)} \tag{1.37}$$

Theorem: If $\mathbf{X} \sim \mathcal{N}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ then

$$\boxed{\mathbf{Y} = \mathbf{AX} + \mathbf{b} \sim \mathcal{N}_n(\mathbf{A}\boldsymbol{\mu} + \mathbf{b})} \tag{1.38}$$

Proof: The proof once again uses the mgf for $\mathbf{Y}$, but also some linear algebra manipulations. □

There are many other theorems and results related to this topic, but I won't go into them for now.

1.8 The t- and F-distributions

These two distributions are useful in certain problems in statistical inference.

1.8.1 The t-distribution

Suppose $W \sim \mathcal{N}(0,1)$ and $V \sim \chi^2(r)$ and that they are independent. Then the joint pdf of W and V, called $h(w,v)$, is the product of the pdf's of W and V:

$$h(w,v) = \begin{cases} \frac{1}{\sqrt{2\pi}} e^{-w^2/2} \frac{1}{\Gamma(r/2)2^{r/2}} v^{r/2-1} e^{-v/2}, & w \in \mathbb{R}, v > 0 \\ 0, & \text{else} \end{cases} \tag{1.39}$$

Now we define a new variable $T = W/\sqrt{V/r}$ and consider the transformation:

$$t = \frac{w}{\sqrt{v/r}} \qquad u = v \tag{1.40}$$

which bijectively maps the parameter space $(w,v) = \mathbb{R} \times \mathbb{R}^+$ to $(t,u) = \mathbb{R} \times \mathbb{R}^+$. The absolute value of the Jacobian of the transformation is given by

$$|J| = \left| \det \begin{pmatrix} \partial_t w & \partial_u w \\ \partial_t v & \partial_u v \end{pmatrix} \right| = \frac{\sqrt{u}}{\sqrt{r}}. \tag{1.41}$$

With this, the joint pdf of T and $U \equiv V$ is given by

$$g(t,u) = |J| h\left(\frac{t\sqrt{u}}{\sqrt{r}}, u\right) = \begin{cases} \frac{u^{r/2-1}}{\sqrt{2\pi}\Gamma(r/2)2^{r/2}} \exp\left[-\frac{u}{2}\left(1 + \frac{t^2}{r}\right)\right] \frac{\sqrt{u}}{\sqrt{r}}, & t \in \mathbb{R}, u \in \mathbb{R}^+ \\ 0, & \text{else} \end{cases} \tag{1.42}$$

By integrating out u we obtain the marginal pdf for T:

$$\begin{aligned} g_1(t) &= \int_{-\infty}^{\infty} g(t,u)\, du \\ &= \int_0^{\infty} \frac{u^{(r+1)/2-1}}{\sqrt{2\pi r}\Gamma(r/2)2^{r/2}} \exp\left[-\frac{u}{2}\left(1 + \frac{t^2}{r}\right)\right] du. \end{aligned} \tag{1.43}$$

Via the substitution $z = u[1 + (t^2/r)]/2$ we can evaluate the integral to find for $t \in \mathbb{R}$

$$\boxed{g_1(t) = \frac{\Gamma[(r+1)/2]}{\sqrt{\pi r}\Gamma(r/2)} \frac{1}{(1+t^2/r)^{(r+1/2)}}} \tag{1.44}$$

A r.v. T with this pdf is said to follow the t-distribution (or the Student's t-distribution) with r degrees of freedom. The t-distribution is symmetric about 0 and has a unique maximum at 0. As $r \in \infty$, the t-distribution converges to $\mathcal{N}(0,1)$.

The mean of $T \sim \mathrm{Stu}(r)$ is zero. The variance can be found to be $\mathrm{Var}(T) = E[T^2] = \frac{r}{r-2}$, so long as $r > 2$.

1.8.2 The F-distribution

Let $U \sim \chi^2(r-1), V \sim \chi^2(r_2)$ be given. Then the joint pdf of U and V is once again the product of their pdf's:

$$h(u,v) = \begin{cases} \frac{1}{\Gamma(r_1/2)\Gamma(r_2/2)2^{(r_1+r_2)/2}} u^{r_1/2-1} v^{r_2/2-1} e^{-(u+v)/2}, & u, v \in \mathbb{R}^+ \\ 0, \quad \text{else} \end{cases} \tag{1.45}$$

Define the new random variable

$$W = \frac{U/r_1}{V/r_2} \tag{1.46}$$

whose pdf $g_1(w)$ we are interested in finding. Consider the transformation

$$w = \frac{u/r_1}{v/r_2}, \quad z = v \tag{1.47}$$

which bijectively maps $(u,v) = \mathbb{R}^+ \times \mathbb{R}^+ \to (w,z) = [\mathbb{R}^+ \times \mathbb{R}^+]$. Like last time, the absolute value of the Jacobian can be found to be

$$|J| = \frac{r_1}{r_2} z. \tag{1.48}$$

The joint pdf $g(w,z)$ of the random variables W and $Z = V$ is obtained from by scaling $h(u,v)$ by $|J|$ and applying the variable transformation:

$$g(w,z) = \frac{1}{\Gamma(r_1/2)\Gamma(r_2/2)2^{\frac{r_1+r_2}{2}}} \left(\frac{r_1 z w}{r_2}\right)^{\frac{r_1-2}{2}} z^{\frac{r_2-2}{2}} \exp\left[-\frac{z}{2}\left(\frac{r_1 w}{r_2}+1\right)\right] \frac{r_1 z}{r_2} \tag{1.49}$$

so long as $(w, z) \in \mathbb{R}^+ \times \mathbb{R}^+$ and 0 otherwise. We then proceed to find the marginal pdf $g_1(w)$ of W by integrating out z. By considering the change of variables:

$$y = \frac{z}{2}\left(\frac{r_1 w}{r_2} + 1\right) \tag{1.50}$$

we can evaluate the integral and find the marginal pdf of W to be

$$g_1(w) = \begin{cases} \frac{\Gamma[(r_1+r_2)/2](r_1/r_2)^{r_1/2}}{\Gamma(r_1/2)\Gamma(r_2/2)} \frac{w^{r_1/2-1}}{(1+r_1 w/r_2)^{(r_1+r_2)/2}}, & w \in \mathbb{R}^+ \\ 0, \quad \text{else} \end{cases} \tag{1.51}$$

W, which is the ratio of two independent chi-quare variables U, V, is said to follow an F-distribution with degrees of freedom r_1 and r_2. We call the ratio $W = (U/r_1)/(V/r_2)$ the "F" ratio.

The mean of W is $E[F] = \frac{r_2}{r_2-2}$. When r_2 is large, $E[F] \to 1$.

1.8.3 The Student's Theorem

Here we will create the connection between the normal distribution and the t-distribution. This is an important result for the later topics on inference for normal random variables.

Theorem: Let $X_1, \ldots, X_n$ be iid r.v. with $X_i \sim \mathcal{N}(\mu, \sigma^2)\forall i$. Define the r.v.'s

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \tag{1.52}$$

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2. \tag{1.53}$$

Then

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

(b) $\bar{X}$ and S^2 are independent.

(c) $(n-1)S^2/\sigma^2 \sim \chi^2(n-1)$.

(d) The variable $\bar{T} = (\bar{X} - \mu)/(S/\sqrt{n})$ follows the Student's t-distribution with $n-1$ degrees of freedom.

$$\bar{T} = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim \text{Stu}(n-1). \tag{1.54}$$

Proof: The proof is good, so I will reproduce it here. Because $X_i \sim \mathcal{N}(\mu, \sigma^2)\forall i$, $\mathbf{X} \sim \mathcal{N}_n\left(\mu\mathbf{1}, \sigma^2\mathbf{1}\right)$, where $\mathbf{1}$ denotes the n-vector whose components are all 1.

Now, consider $\mathbf{v}^\top = (1/n)\mathbf{1}^\top$. We see that $\bar{X} = \mathbf{v}^\top\mathbf{X}$. Define the random vector $\mathbf{Y} = (X_1 - \bar{X}, \ldots, X_2 - \bar{X})^\top$ and consider the (true) equality:

$$\mathbf{W} = \begin{pmatrix} \bar{X} \\ \mathbf{Y} \end{pmatrix} = \underbrace{\begin{pmatrix} \mathbf{v}^\top \\ \mathbb{I} - \mathbf{1}\mathbf{v}^\top \end{pmatrix}}_{\text{the transformation}} \mathbf{X} \tag{1.55}$$

which just restates our definitions nicely. We see that $\mathbf{W}$ is a result of a linear transformation of multivariate normal random vector, and so it follows that $\mathbf{W} \sim \mathcal{N}_{n+1}$ with mean

$$E[\mathbf{W}] = \begin{pmatrix} \mathbf{v}^\top \\ \mathbb{I} - \mathbf{1}\mathbf{v}^\top \end{pmatrix} \mu\mathbf{1} = \begin{pmatrix} \mu \\ \mathbf{0}_n \end{pmatrix} \tag{1.56}$$

and the covariance matrix

$$\boldsymbol{\Sigma} = \begin{pmatrix} \mathbf{v}^\top \\ \mathbb{I} - \mathbf{1}\mathbf{v}^\top \end{pmatrix} \sigma^2\mathbb{I} \begin{pmatrix} \mathbf{v}^\top \\ \mathbb{I} - \mathbf{1}\mathbf{v}^\top \end{pmatrix}^\top = \sigma^2 \begin{pmatrix} \frac{1}{n} & \mathbf{0}_n^\top \\ \mathbf{0}_n & \mathbb{I} - \mathbf{1}\mathbf{v}^\top \end{pmatrix} \tag{1.57}$$

From here, part (a) is proven. Next, observe that $\boldsymbol{\Sigma}$ is diagonal, and so all covariances are zero. This means $\bar{X}$ is independent of $\mathbf{Y}$. But because $S^2 = (n-1)^{-1}\mathbf{Y}^\top\mathbf{Y}$, $\bar{X}$ is independent of S^2 as well. So, (b) is proven.

Now, consider the r.v.

$$V = \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma} \right)^2 \tag{1.58}$$

Each summand of V is a square of an $\mathcal{N}(0,1)$ r.v., and so each follows a $\chi^2(1)$. Because V is a sum of squares of n such $\chi^2(1)$'s, $V \sim \chi^2(n)$. Next, we can rewrite V as

$$V = \sum_{i=1}^{n} \left(\frac{(X_i - \bar{X}) + (\bar{X} - \mu)}{\sigma} \right)^2 = \frac{(n-1)S^2}{\sigma^2} + \left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \right)^2. \tag{1.59}$$

By (b), the summands in the last equation is are independent. The second term is a square of a $\mathcal{N}(0,1)$, so it follows a $\chi^2(1)$. Taking mgfs of both sides, we get

$$(1-2t)^{-n/2} = \underbrace{E[\exp\{t(n-1)S^2/\sigma^2\}]}_{M_{(c)}}(1-2t)^{-1/2}. \tag{1.60}$$

Solving for the mgf of $(n-1)S^2/\sigma^2$ we get part (c). Finally, writing T as

$$T = \frac{(\bar{X} - \mu)/(\sigma/\sqrt{n})}{\sqrt{(n-1)S^2/(\sigma^2(n-1))}} \tag{1.61}$$

and using (a)-(c) gives us (d). *Hint*: consider what distributions the numerator and denominator of T follow. □

1.9 Problems

3.6.4

(a) X has a standard normal distribution:

```
x=seq(-6,6,.01); plot(dnorm(x)~x)
```

(b) X has a t-distribution with 1 degree of freedom.

```
lines(dt(x,1)~x,lty=2)
```

(c) X has a t-distribution with 3 degrees of freedom.

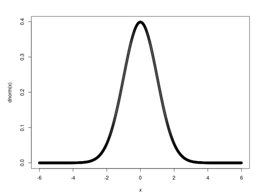
```
lines(dt(x,3)~x,lty=2)
```

(d) X has a t-distribution with 10 degrees of freedom.

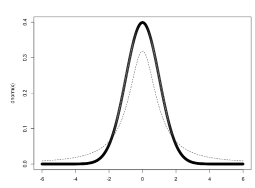
```
lines(dt(x,10)~x,lty=2)
```

(e) X has a t-distribution with 30 degrees of freedom.

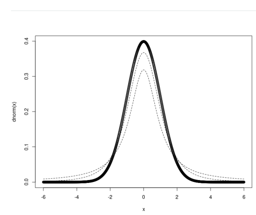
```
lines(dt(x,30)~x,lty=2)
```



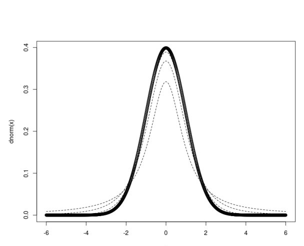
(a)



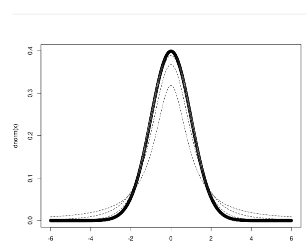
(b)



(c)



(d)



(e)

3.6.5

(a) $P(|X| \geq 2) = 2 \times [1 - P(X \leq 2)] = \mathbf{0.046}$.

```
> 2*(1 - pnorm(2))
[1] 0.04550026
```

(b) $P(|X| \geq 2) = 2 \times [1 - P(X \leq 2)] = \mathbf{0.295}$.

```
> 2*(1 - pt(2,1))
[1] 0.2951672
```

(c) $P(|X| \geq 2) = 2 \times [1 - P(X \leq 2)] = \mathbf{0.139}$.

```
> 2*(1 - pt(2,3))
[1] 0.139326
```

(d) $P(|X| \geq 2) = 2 \times [1 - P(X \leq 2)] = \mathbf{0.073}$.

```
> 2*(1 - pt(2,10))
[1] 0.07338803
```

(e) $P(|X| \geq 2) = 2 \times [1 - P(X \leq 2)] = \mathbf{0.055}$.

```
> 2*(1 - pt(2,30))
[1] 0.05462504
```

3.6.11: Let $T = W/\sqrt{V/r}$, where the independent variables $W \sim \mathcal{N}(0,1)$ and $V \sim \chi^2(r)$. Show that $T^2 \sim F(r_1 = 1, r_2 = r)$. *Hint*: What is the distribution of the numerator of T^2?

Solution: Let the independent random variables U, V be given, with $W \sim \mathcal{N}(0,1)$ and $U \sim \chi^2(r)$. The random variable T^2, where $T = W/\sqrt{V/r}$ is given by

$$T^2 = \left(\frac{W}{\sqrt{V/r}}\right)^2 = \frac{W^2}{V/r}. \tag{1.62}$$

Because $W \sim \mathcal{N}(0,1)$, we have that $W^2 \sim \chi^2(1)$ (by theorem). Now, T^2 has the form

$$T^2 = \frac{W^2}{V/r} = \frac{W^2/1}{V/r} \tag{1.63}$$

where 1 is the df of $\chi^2(1)$ which W follows, and r is the df of $\chi^2(r)$ which U follows. Thus, $T^2 \sim F(1,r)$, by the definition of the F-distribution. $\square$

3.6.15: Let X_1, X_2 be iid with common distribution having the pdf

$$f(x) = \begin{cases} e^{-x}, & 0 < x < \infty \\ 0, & \text{else} \end{cases} \tag{1.64}$$

Show that $Z = X_1/X_2$ has an F-distribution.

Solution: It suffices to show that Z can be written as a ratio of two χ^2-distributed independent random variables. To this end, we can consider the mgf $M_X(t)$ of X_1, which is also identically that of X_2 since X_1, X_2 are iid:

$$M_X(t) = E[e^{tx}] = \int_0^\infty e^{tx}e^{-x}\,dx = (1-t)^{-1}. \tag{1.65}$$

However, this does not quite match the mgf for a $\chi^2(2)$. To circumvent this problem, we rewrite

$$Z = \frac{X_1}{X_2} = \frac{2X_1/2}{2X_2/2} = \frac{(X_1+X_1)/2}{(X_2+X_2)/2}, \tag{1.66}$$

as we expect $r = 2$. Let $Y_1 = X_1 + X_1$. Then we have trivially $Y_1 = 2X_1$, and so $|J| = 2$. With this, Y_1 has the pdf

$$\tilde{f}_Y(y) = |J|f(x) = 2f(x) = \begin{cases} 2e^{-y/2}, & 0 < y < \infty \\ 0, & \text{else} \end{cases}. \tag{1.67}$$

From here, we find the mgf of Y_1 to be

$$M_{Y_1}(t) = E[e^{ty}] = \frac{1}{2}\int_0^\infty e^{ty}e^{-y/2}\,dy = (1-2t)^{-1} = (1-2t)^{-2/2}, t < \frac{1}{2}. \tag{1.68}$$

By symmetry, $M_{Y_2}(t)$ is identically $M_{Y_1}(t)$, and both are the mgf for $\chi^2(r=2)$. Because each mgf uniquely determines a pdf, $Y_1, Y_2 \sim \chi^2(r=2)$ identically and independently (for each depends exclusively on X_1, X_2, respectively). Therefore,

$$Z = \frac{(X_1+X_1)/2}{(X_2+X_2)/2} = \frac{Y_1/2}{Y_2/2} \tag{1.69}$$

follows the F-distribution with degrees of freedom $r_1 = r_2 = 2$, by definition. $\square$

3.6.16: Let X_1, X_2, X_3 be independent r.v. with $X_i \sim \chi^2(r_i)$.

(a) Show that $Y_1 = X_1/X_2$ and $Y_2 = X_1 + X_2$ are independent and that $Y_2 \sim \chi^2(r_1 + r_2)$.

(b) Deduce that

$$\frac{X_1/r_1}{X_2/r_2} \text{ and } \frac{X_3/r_3}{(X_1+X_2)/(r_1+r_2)} \tag{1.70}$$

are independent F-variables.

Solution:

(a) We consider the transformation

$$y_1 = u(x_1, x_2) = \frac{x_1}{x_2} \tag{1.71}$$

$$y_2 = v(x_1, x_2) = x_1 + x_2. \tag{1.72}$$

whose inverse is

$$\begin{aligned} x_1 &= \bar{u}(y_1, y_2) = \frac{y_1 y_2}{1 + y_1} \\ x_2 &= \bar{v}(y_1, y_2) = \frac{y_2}{1 + y_1}. \end{aligned} \tag{1.73}$$

The absolute value of the Jacobian is

$$|J| = \left| \det \begin{pmatrix} \partial_{y_1} \bar{u} & \partial_{y_2} \bar{u} \\ \partial_{y_1} \bar{v} & \partial_{y_2} \bar{v} \end{pmatrix} \right| = \frac{y_2}{(1+y_1)^2}, \tag{1.74}$$

which maps one-to-one from the space of X_1, X_2 $\mathbb{R}^+ \times \mathbb{R}^+$ onto the space of Y_1, Y_2 $\mathbb{R}^+ \times \mathbb{R}^+$. Since X_1, X_2 are independent, we consider the joint pdf of X_1, X_2:

$$h(x_1, x_2) = \begin{cases} \frac{x_1^{r_1/2-1} x_2^{r_2/2-1}}{\Gamma(r_1/2)\Gamma(r_2/2)2^{(r_1+r_2)/2}} e^{-(x_1+x_2)/2}, & 0 < x_1, x_2 < \infty \\ 0, & \text{else} \end{cases} \tag{1.75}$$

from which we can deduce the joint pdf for Y_1, Y_2:

$$\begin{aligned} \tilde{h}(y_1, y_2) &= |J| h\left(\frac{y_1 y_2}{1+y_1}, \frac{y_2}{1+y_1}\right) \\ &= \begin{cases} \frac{y_2 (y_1 y_2)^{r_1/2-1} y_2^{r_2/2-1} (1+y_1)^{-r_1/2-r_2/2+\cancel{2}}}{\cancel{(1+y_1)^2}\,\Gamma(r_1/2)\Gamma(r_2/2)2^{(r_1+r_2)/2}} e^{-y_2/2}, & 0 < y_1, y_2 < \infty \\ 0, & \text{else} \end{cases} \\ &= \begin{cases} \frac{y_2^{r_1/2+r_2/2-1} y_1^{r_1/2-1} (1+y_1)^{-r_1/2-r_2/2}}{\Gamma(r_1/2)\Gamma(r_2/2)2^{(r_1+r_2)/2}} e^{-y_2/2}, & 0 < y_1, y_2 < \infty \\ 0, & \text{else} \end{cases} \end{aligned} \tag{1.76}$$

Without further computation we see that $\tilde{h}(y_1, y_2)$ can be written as a product of two nonnegative functions of y_1 and y_2. In view of Theorem 2.4.1, Y_1 and Y_2 are independent. □

Next, we wish to show $Y_2 \sim \chi^2(X_1, X_2)$, to which end we find the marginal pdf $g_2(y_2)$ of Y_2:

$$\begin{aligned} g_2(y_2) &= \int_0^\infty \tilde{h}(y_1, y_2)\,dy_1 \\ &= \mathfrak{C}\int_0^\infty y_1^{r_1/2-1}(1+y_1)^{-r_1/2-r_2/2}\,dy_1 \\ &= \mathfrak{C}\frac{\Gamma(r_1/2)\Gamma(r_2/2)}{\Gamma[(r_1+r_2)/2]} \end{aligned} \tag{1.77}$$

where $\mathfrak{C}$ contains all the y_1-independent elements. From here, via simple back-substitution we obtain the marginal pdf for Y_2:

$$g_2(y_2) = \begin{cases} \frac{y_2^{(r_1+r_2)/2-1}}{\Gamma[(r_1+r_2)/2]2^{(r_1+r_2)/2}} e^{-y_2/2}, & 0 < y_2 < \infty \\ 0, & \text{else} \end{cases}, \tag{1.78}$$

i.e., $Y_2 \sim \chi^2(r_1 + r_2)$. □

Mathematica code:

```
In[20]:= Integrate[
x^(r1/2 - 1) (1 + x)^(-r1/2 - r2/2), {x, 0, Infinity}]

Out[20]= ConditionalExpression[(Gamma[r1/2] Gamma[r2/2])/
Gamma[(r1 + r2)/2], Re[r2] > 0 && Re[r1] > 0]
```

(b) By definition, because X_1, X_2 are independent random variables with $X_i \sim \chi^2(r_i)$,

$$\Omega = \frac{X_1/r_1}{X_2/r_2} \sim F(r_1, r_2). \tag{1.79}$$

Also, because $X_3 \sim \chi^2(r_3)$ and $(X_1 + X_2) \sim \chi^2(r_1 + r_2)$ (from (a)), we have

$$\Lambda = \frac{X_3/r_3}{(X_1+X_2)/(r_1+r_2)} \sim F(r_3, r_1 + r_2) \tag{1.80}$$

as well. Furthermore, because

$$\Omega = \frac{X_1/r_1}{X_2/r_2} = \frac{r_2}{r_1}Y_1 \tag{1.81}$$

$$\Lambda = \frac{r_1+r_2}{r_3}\frac{X_3}{Y_2} \tag{1.82}$$

and because X_1, X_2, X_3 are independent, we have that Y_1, Y_2, X_3 are independent. Therefore, it is necessary that $\Omega \sim F(r_1, r_2)$ and $\Lambda \sim F(r_3, r_1 + r_2)$ are independent as well. □

Part 2

Elementary Statistical Inferences

2.1 Sampling & Statistics

2.1.1 Point estimators

2.1.2 Histogram estimates of pmfs and pdfs

The distribution of X is discrete

The distribution of X is continuous

2.2 Confidence Intervals

2.2.1 CI for difference in means

2.2.2 CI for difference in proportions

2.3 Order Statistics

2.3.1 Quantiles

2.3.2 CI for quantiles

2.4 Problems

4.1.1
4.1.3
4.1.8