1. A study compared frequencies of a particular allele in a sample of adult-onset diabetics and a sample of non-diabetics. Here are the data:

	Diabetic	Normal
Bb or bb	12	4
BB	39	49

Test the null hypothesis that BB occurs just as frequently among diabetics as among normal persons in the population from which this sample was taken, at the 5% level.

Answer: We need the marginal totals:

	Diabetic	Normal	
Bb or bb	12	4	16
BB	39	49	88
	51	53	104

We take $\dfrac{16}{104}$ to be an estimate of an individual's probability of being in the first row of the table, and similarly we take $\dfrac{51}{104}$ to be an estimate of an individual's probability of being in the first column. The null hypothesis says the events of being in the first row and being in the first column are independent, and so an estimate of the probability of being in the first row and the first column is $\dfrac{16}{104} \cdot \dfrac{51}{104}$. As an estimate of the expected number of individuals in the first row and first column, we get

$$104 \cdot \frac{16}{104} \cdot \frac{51}{104} = \frac{16 \cdot 51}{104}.$$

In the sum $\sum \dfrac{(\text{observed} - \text{expected})^2}{\text{expected}}$, the term corresponding to the first row and first column is therefore $\dfrac{(12 - (16 \cdot 51/104))^2}{16 \cdot 51/104}$. The test statistic is the sum of the terms similarly corresponding to the four cells; it is

$$\frac{(12 - (16 \cdot 51/104))^2}{16 \cdot 51/104} + \frac{(4 - (16 \cdot 53/104))^2}{16 \cdot 53/104} + \frac{(39 - (88 \cdot 51/104))^2}{16 \cdot 51/104} + \frac{(49 - (88 \cdot 53/104))^2}{16 \cdot 53/104}$$

$$= 5.099788\ldots$$

Under the null hypothesis this test statistic has a $\chi^2_{(2-1)(2-1)} = \chi^2_1$ distribution. According to the table, $\mathbf{Pr}(\chi^2_1 > 3.841) = 0.95$. Since $5.099788 > 3.841$, we reject the null hypothesis.

2. Suppose $Y_1, \ldots, Y_n$ are independent and $Y_i \sim N(\alpha_0 + \alpha_1 x_i, \sigma^2)$ for $i = 1, \ldots, 8$. Find the probability distribution of $\sum_{i=1}^{8}(Y_i - (\widehat{\alpha}_0 + \widehat{\alpha}_1 x_i))^2$, where $\widehat{\alpha}_0$ and $\widehat{\alpha}_1$ are the respective least-squares estimates of α_0 and α_1.

<u>**Answer:**</u> I intended to write $Y_1, \ldots, Y_8$ above, and that is how apparently everyone read it. If the least-squares estimates $\widehat{\alpha}_0$ and $\widehat{\alpha}_1$ were based on n observations for some $n > 8$ then we would have a more complicated problem.

Let $Y = \begin{bmatrix} Y_1 \\ \vdots \\ Y_8 \end{bmatrix}$, $X = \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_8 \end{bmatrix}$, and $\alpha = \begin{bmatrix} \alpha_0 \\ \alpha_1 \end{bmatrix}$. Then we have $Y \sim N_8(X\alpha, \sigma^2 I_8)$. The least-squares estimators are given by $\widehat{\alpha} = \begin{bmatrix} \widehat{\alpha}_0 \\ \widehat{\alpha}_1 \end{bmatrix} = (X'X)^{-1}X'Y$ and the fitted values by

$$\widehat{Y} = \begin{bmatrix} \widehat{\alpha}_0 + \widehat{\alpha}_1 x_1 \\ \vdots \\ \widehat{\alpha}_0 + \widehat{\alpha}_1 x_8 \end{bmatrix} = \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_8 \end{bmatrix} \begin{bmatrix} \widehat{\alpha}_0 \\ \widehat{\alpha}_1 \end{bmatrix} = X\widehat{\alpha} = X(X'X)^{-1}X'Y = HY.$$

The matrix $H = X(X'X)^{-1}X'$ is a symmetric idempotent matrix whose column space is the same as that of X, and whose rank is therefore 2 (assuming the x_is are not all equal). We want the distribution of

$$\sum_{i=1}^{8}(Y_i - (\widehat{\alpha}_0 + \widehat{\alpha}_1 x_i))^2 = \|Y - \widehat{Y}\|^2 = \|Y - HY\|^2 = \|(I - H)Y\|^2.$$

Observe that

$$\mathbf{E}((I - H)Y) = (I - H)\,\mathbf{E}(Y) = (I - H)X\alpha = X\alpha - HX\alpha = X\alpha - X\alpha = 0.$$

Since $I - H$ is a symmetric idempotent matrix of rank $8 - 2 = 6$ whose columns are orthogonal to the columns of X, we have

$$I - H = G \begin{bmatrix} 0 & & \\ & 0 & \\ & & I_{8-2} \end{bmatrix} G'$$

where $G = [g_1, g_2, g_3, \ldots, g_8] \in \mathbb{R}^8$ is an orthogonal matrix whose first two columns span the column space of X. So we have

$$\begin{aligned} \|(I - H)Y\|^2 &= Y'(I - H)'(I - H)Y = Y'(I - H)Y \\ &= Y'G \begin{bmatrix} 0 & & \\ & 0 & \\ & & I_{8-2} \end{bmatrix} G'Y = W' \begin{bmatrix} 0 & & \\ & 0 & \\ & & I_{8-2} \end{bmatrix} W = W_3^2 + \cdots + W_8^2. \end{aligned}$$

Now

$$W = G'Y \sim N_8(G'X\alpha,\, \sigma^2 G' I_8 G) = N_8(G'X\alpha,\, \sigma^2 I_8)$$

The expected value $G'X\alpha$ is CONTINUED⟶

$$G'X\alpha = \begin{bmatrix} g_1'X\alpha \\ g_2'X\alpha \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

i.e., the 3rd through 8th entries are 0 because $g_3, \ldots, g_8$ are orthogonal to the column space of X. Consequently

$$\begin{bmatrix} W_3 \\ \vdots \\ W_8 \end{bmatrix} \sim N_{8-2}(0,\, \sigma^2 I_{8-2}),$$

or, in other words,

$$W_3, \ldots, W_8 \sim \text{i.i.d.}\, N_1(0, \sigma^2)$$

and so

$$\sum_{i=1}^{8} (Y_i - (\widehat{\alpha}_0 + \widehat{\alpha}_1 x_i))^2 = \|(I-H)Y\|^2 = W_3^2 + \cdots + W_8^2 \sim \sigma^2\chi^2_{8-2} = \sigma^2\chi^2_6.$$

And that answers the question.

Why did I mention the fact that $\mathbf{E}((I-H)Y) = 0$ above? Because you could rely on that as an alternative to mentioning that all but the first two components of the vector $G'X\alpha$ are 0, although we never emphasized that point.

It is amusing, and consequently perhaps useful as an aid to memory, to notice that the practice of calling the diacritical mark above the "Y" in "$\widehat{Y}$" a "hat", conjoined with the identity $HY = \widehat{Y}$, may be the cause of the convention of using the letter H for that matrix, and sometimes even calling it "the hat matrix."

3. Suppose $Y_{ij} \sim N(\mu_i,\, \sigma^2)$ for $i = 1, 2, 3$ and $j = 1, 2, 3, 4$, and all 12 of the Ys are mutually independent. The entry in the i^{th} column and j^{th} row below is the observed value of Y_{ij}. Test the null hypothesis $\mu_1 = \mu_2 = \mu_3$ at the 5% level, making it clear what you do and how you do it.

29	20	17
19	19	16
34	26	7
32	31	28

Answer: This is a standard analysis-of-variance problem.

Total variability in the data, with $3 \cdot 4 - 1 = 11$ degrees of freedom

Between-group variability, with $3 - 1 = 2$ degrees of freedom

Within-group variability, with $3 \cdot (4 - 1) = 9$ degrees of freedom

$$\underbrace{\sum_{i=1}^{3}\sum_{j=1}^{4} \left(Y_{ij} - \overline{Y}_{\bullet\bullet}\right)^2}_{\text{Total}} = \underbrace{\sum_{i=1}^{3}\sum_{j=1}^{4} \left(Y_{i\bullet} - \overline{Y}_{\bullet\bullet}\right)^2}_{\text{Between}} + \underbrace{\sum_{i=1}^{3}\sum_{j=1}^{4} \left(Y_{ij} - \overline{Y}_{i\bullet}\right)^2}_{\text{Within}}.$$

Notice that

- The between-group sum of squares is distributed as $\sigma^2\chi^2_2$ <u>if</u> the null hypothesis is true, and
- The within-group sum of squares is distributed as $\sigma^2\chi^2_9$ (regardless of whether the null hypothesis is true), and
- They are independent, and so
- The former divided by the latter is distributed as $F_{2,9}$ if the null hypothesis is true. (The "nuisance parameter" σ^2 cancels out when we divide.)

According to the table, $\mathbf{Pr}(F_{2,9} > 4.26) = 0.05$, so we reject the null hypothesis iff $F > 4.26$.

From the data above, we get

$$\begin{aligned}
\overline{Y}_{1\bullet} &= (29+19+34+32)/4 &= 28.5. \\
\overline{Y}_{2\bullet} &= (20+19+26+31)/4 &= 24. \\
\overline{Y}_{3\bullet} &= (17+16+7+28)/4 &= 17. \\
\overline{Y}_{\bullet\bullet} &= 23+1/6 \longleftarrow \text{ saving rounding for the last step}
\end{aligned}$$

$$\begin{aligned}
4\left(\overline{Y}_{1\bullet}-\overline{Y}_{\bullet\bullet}\right)^2 + 4\left(\overline{Y}_{1\bullet}-\overline{Y}_{\bullet\bullet}\right)^2 + 4\left(\overline{Y}_{1\bullet}-\overline{Y}_{\bullet\bullet}\right)^2 &= 268+2/3 \\
&= \text{between-group sum of squares.}
\end{aligned}$$

$$\begin{aligned}
\left(Y_{11}-\overline{Y}_{1\bullet}\right)^2 + \left(Y_{12}-\overline{Y}_{1\bullet}\right)^2 + \left(Y_{13}-\overline{Y}_{1\bullet}\right)^2 &= (29-28.5)^2+(19-28.5)^2 \\
&\quad +(34-28.5)^2+(32-28.5)^2 = 133.
\end{aligned}$$

$$\begin{aligned}
\left(Y_{21}-\overline{Y}_{2\bullet}\right)^2 + \left(Y_{22}-\overline{Y}_{2\bullet}\right)^2 + \left(Y_{23}-\overline{Y}_{2\bullet}\right)^2 &= (20-24)^2+(19-24)^2 \\
&\quad +(26-24)^2+(31-24)^2 = 94.
\end{aligned}$$

$$\begin{aligned}
\left(Y_{31}-\overline{Y}_{3\bullet}\right)^2 + \left(Y_{32}-\overline{Y}_{3\bullet}\right)^2 + \left(Y_{33}-\overline{Y}_{3\bullet}\right)^2 &= (17-17)^2+(16-17)^2 \\
&\quad +(7-17)^2+(28-17)^2 = 222.
\end{aligned}$$

$$\begin{aligned}
133+94+222 &= 449 \\
&= \text{between-group sum of squares}
\end{aligned}$$

$$F = \frac{\text{between-group sum of squares/degrees of freedom}}{\text{within-group sum of squares/degrees of freedom}} = \frac{(268+2/3)/2}{449/9} = 2.69265\ldots \not> 4.26$$

so we do not reject the null hypothesis.

4. Discard the third column of data above. Find a 90% confidence interval for $\mu_1 - \mu_2$.

<u>**Answer:**</u>

$$\overline{Y}_{1\bullet} = \frac{Y_{11} + Y_{12} + Y_{13} + Y_{14}}{4} \sim N\left(\mu_1, \frac{\sigma^2}{4}\right)$$

$$\overline{Y}_{2\bullet} = \frac{Y_{21} + Y_{22} + Y_{23} + Y_{24}}{4} \sim N\left(\mu_2, \frac{\sigma^2}{4}\right)$$

$$\overline{Y}_{1\bullet} - \overline{Y}_{2\bullet} \sim N\left(\mu_1 - \mu_2, \frac{\sigma^2}{4} + \frac{\sigma^2}{4}\right) = N\left(\mu_1 - \mu_2, \frac{\sigma^2}{2}\right)$$

$$\frac{\left(\overline{Y}_{1\bullet} - \overline{Y}_{2\bullet}\right) - (\mu_1 - \mu_2)}{\sigma/\sqrt{2}} \sim N(0,1)$$

$$\frac{\sum_{j=1}^4 \left(Y_{1j} - \overline{Y}_{1\bullet}\right)^2 + \sum_{j=1}^4 \left(Y_{2j} - \overline{Y}_{2\bullet}\right)^2}{\sigma^2} \sim \chi^2_{(4-1)+(4-1)} = \chi^2_6$$

and the last two random variables above are independent. Therefore

$$\frac{\sqrt{2}\left(\left(\overline{Y}_{1\bullet} - \overline{Y}_{2\bullet}\right) - (\mu_1 - \mu_2)\right)}{\sqrt{\left(\sum_{j=1}^4 \left(Y_{1j} - \overline{Y}_{1\bullet}\right)^2 + \sum_{j=1}^4 \left(Y_{2j} - \overline{Y}_{2\bullet}\right)^2\right)/6}} \sim t_6. \qquad \text{(“}\sigma\text{” has cancelled out.)}$$

The table then tells us that

$$\mathbf{Pr}\left(-1.943 < \frac{\sqrt{2}\sqrt{6}\left(\left(\overline{Y}_{1\bullet} - \overline{Y}_{2\bullet}\right) - (\mu_1 - \mu_2)\right)}{\sqrt{\sum_{j=1}^4 \left(Y_{1j} - \overline{Y}_{1\bullet}\right)^2 + \sum_{j=1}^4 \left(Y_{2j} - \overline{Y}_{2\bullet}\right)^2}} < 1.943\right) = 0.9.$$

Solving the system of two inequalities that appear within the expression "$\mathbf{Pr}(\cdots\cdots)$" above, we find that the endpoints of the 90% confidence interval are:

$$\overline{Y}_{1\bullet} - \overline{Y}_{2\bullet} \pm 1.943\sqrt{\frac{\sum_{j=1}^4 \left(Y_{1j} - \overline{Y}_{1\bullet}\right)^2 + \sum_{j=1}^4 \left(Y_{2j} - \overline{Y}_{2\bullet}\right)^2}{12}}.$$

Then we can plug in the numbers, which are among those on page 4 of these answers:

$$28.5 - 24 \pm 1.943\sqrt{\frac{133 + 94}{12}},$$

and so the 90% confidence interval is the interval from $-3.95\ldots$ to $12.95\ldots$.

5. Suppose $x_1 = 200$, $x_2 = 220$ and $x_3 = 222$, and logit $\mathbf{Pr}(Y_i = 1) = (1/7)(x_i - 221)$ for $i = 1, 2, 3$, and Y_1, Y_2, Y_3 are mutually independent. Find $\mathbf{Pr}(Y_1 = 0 \ \&\ Y_2 = 1 \ \&\ Y_3 = 0)$.

<u>**Answer:**</u> If $v = \operatorname{logit}(p) = \log \dfrac{p}{1-p} = \log\left(-1 + \dfrac{1}{1-p}\right)$ then $p = \dfrac{1}{1+e^{-v}}$ and $1 - p = \dfrac{1}{1+e^{+v}}$. So

$$\mathbf{Pr}(Y_1 = 0 \ \&\ Y_2 = 1 \ \&\ Y_3 = 0) \overset{\boxed{\text{By independence}}}{\downarrow\!=} \mathbf{Pr}(Y_1 = 0)\,\mathbf{Pr}(Y_2 = 1)\,\mathbf{Pr}(Y_3 = 0)$$

$$= (1 - \mathbf{Pr}(Y_1 = 1))\,\mathbf{Pr}(Y_2 = 1)(1 - \mathbf{Pr}(Y_3 = 1))$$

$$= \frac{1}{1 + e^{+(1/7)(200-221)}} \cdot \frac{1}{1 + e^{-(1/7)(220-221)}} \cdot \frac{1}{1 + e^{+(1/7)(222-221)}}$$

$$= \frac{1}{1 + e^{-3}} \cdot \frac{1}{1 + e^{1/7}} \cdot \frac{1}{1 + e^{1/7}} = 0.20539\ldots.$$

6. Suppose $X_1, \ldots, X_n \sim \text{i.i.d.}\, N(\mu, 1^2)$.

(a) Show that the likelihood-ratio test of the null hypothesis $\mu = 3$ against the alternative hypothesis $\mu = 5$ rejects the null hypothesis if and only if $\overline{X} = (X_1 + \cdots + X_n)/n > c$ for some number c.

<u>**Answer:**</u> The likelihood function is

$$L(\mu) = L(\mu \mid X_1 = x_1, \ldots, X_n = x_n) = \prod_{i=1}^{n}\left(\frac{1}{\sqrt{2\pi}}\exp\left(\frac{-(x_i-\mu)^2}{2}\right)\right)$$

$$= (2\pi)^{-n/2}\exp\left(\frac{-\sum_{i=1}^{n}(x_i-\mu)^2}{2}\right) = (2\pi)^{-n/2}\exp\left(\frac{-n(\overline{x}-\mu)^2 - \sum_{i=1}^{n}(x_i-\overline{x})^2}{2}\right).$$

Therefore the likelihood ratio for this particular pair of hypotheses is:

$$\frac{L(3 \mid \overline{X})}{L(5 \mid \overline{X})} = \frac{\exp\left(\dfrac{-n(\overline{X}-3)^2 - \sum_{i=1}^{n}(X_i-\overline{X})^2}{2}\right)}{\exp\left(\dfrac{-n(\overline{X}-5)^2 - \sum_{i=1}^{n}(X_i-\overline{X})^2}{2}\right)} = \exp\left(\frac{n}{2}\left\{(\overline{X}-5)^2 + (\overline{X}-3)^2\right\}\right)$$

$$= \exp\left(\frac{n}{2}\left(34 - 16\overline{X}\right)\right), \text{ and this is a decreasing function of } \overline{X}.$$

We reject the null hypothesis iff this likelihood ratio is too small. Since this likelihood ratio is a decreasing function of $\overline{X}$, we reject the null hypothesis iff $\overline{X}$ is too big.

(b) Find the value of c if $n = 6$ and the probability of Type I error is 0.05.

Answer: We have $\overline{X} \mid [\mu = 3] \sim N(3,\, 1^2/6)$, so

$$0.05 = \mathbf{Pr}\left(\overline{X} > c \mid \mu = 3\right) = \mathbf{Pr}\left(\frac{\overline{X}-3}{1/\sqrt{6}} > \frac{c-3}{1/\sqrt{6}} \,\middle|\, \mu = 3\right) = \mathbf{Pr}(Z > \sqrt{6}(c-3)) = 1-\Phi(\sqrt{6}(c-3)).$$

$$\begin{aligned} \Phi(\sqrt{6}(c-3)) &= 0.95. \\ \sqrt{6}(c-3) &= \Phi^{-1}(0.95) = 1.645 \text{ (from the table)}. \\ c &= 3 + \frac{1.645}{\sqrt{6}} = 3.67\ldots. \end{aligned}$$

7. The frequency P with which a possibly biased coin turns up HEADS is distributed as $\text{Beta}(3,3)$. Let X be the number of times the coin must be tossed in order to get HEADS once. Describe the posterior distribution of P given the observation that $X = 3$.

Answer: Since $P \sim \text{Beta}(3,3)$, the prior density is

$$f_P(p) = [\text{constant}] \cdot p^{3-1}(1-p)^{3-1}.$$

The likelihood function is

$$L(p) = \mathbf{Pr}(X = 3 \mid P = p) = \mathbf{Pr}(\text{TAILS on 1st \& 2nd trials \& HEADS on 3rd} \mid P = p) = (1-p)^2 p.$$

The posterior density is therefore

$$[\text{constant}] \cdot p^{3-1}(1-p)^{3-1} \cdot (1-p)^2 p = [\text{constant}] \cdot p^{4-1}(1-p)^{5-1}.$$

In short, $P \mid [X = 3] \sim \text{Beta}(4,5)$.

8. Recall that the Fisher information is $I(\theta) = \mathbf{var}_\theta\left((d/d\theta)\log f(X \mid \theta)\right)$. Suppose $X \sim \text{Bin}(3,\theta)$. Use the Fisher information to find a lower bound on the mean squared error of unbiased estimators of θ.

Answer: The Cramér-Rao lower bound is $1/I(\theta)$. First, we find $I(\theta)$:

$$\begin{aligned} I(\theta) &= \mathbf{var}_\theta\left((d/d\theta)\log f(X \mid \theta)\right) = \mathbf{var}_\theta\left(\frac{d}{d\theta}\log\left\{\binom{3}{X}\theta^X(1-\theta)^{3-X}\right\}\right) \\ &= \mathbf{var}_\theta\left(\frac{d}{d\theta}\left\{\log\binom{3}{X} + X\log\theta + (3-X)\log(1-\theta)\right\}\right) = \mathbf{var}_\theta\left(\frac{X}{\theta} - \frac{3-X}{1-\theta}\right) \\ &= \mathbf{var}_\theta\left(\frac{X-3\theta}{\theta(1-\theta)}\right) = \frac{\mathbf{var}_\theta(X-3\theta)}{\theta^2(1-\theta)^2} = \frac{\mathbf{var}_\theta(X)}{\theta^2(1-\theta)^2} = \frac{3\theta(1-\theta)}{\theta^2(1-\theta)^2} = \frac{3}{\theta(1-\theta)}. \end{aligned}$$

So the lower bound is $\dfrac{\theta(1-\theta)}{3}$. In other words, no unbiased estimator of θ based on the observed value of this random variable X can have a mean squared error smaller than $\dfrac{\theta(1-\theta)}{3}$.

9. Suppose $X_1, X_2 \sim$ i. i. d. and $X_1 = \left\{ \begin{array}{ll} 1 & \text{with probability } \theta \\ 0 & \text{with probability } 1-\theta \end{array} \right\}$.

(a) Show that X_1 is an unbiased estimator of θ.

Answer: $\mathbf{E}_\theta(X_1) = 0 \cdot \mathbf{Pr}_\theta(X_1 = 0) + 1 \cdot \mathbf{Pr}_\theta(X_1 = 1) = \mathbf{Pr}(X_1 = 1) = \theta$.

(b) Show that $X_1 + X_2$ is sufficient for θ.

Answer: One way to do this is by finding a Fisher factorization. Notice that

$$\mathbf{Pr}_\theta(X_1 = x_1) = \left\{ \begin{array}{ll} \theta & \text{if } x_1 = 1 \\ 1-\theta & \text{if } x_1 = 0 \end{array} \right\} = \theta^{x_1}(1-\theta)^{1-x_1}.$$

Consequently

$$\begin{aligned} f_{X_1,X_2}(x_1, x_2 \mid \theta) &= \mathbf{Pr}_\theta(X_1 = x_1 \ \& \ X_2 = x_2) = \mathbf{Pr}_\theta(X_1 = x_1)\, \mathbf{Pr}_\theta(X_2 = x_2) \\ &= \theta^{x_1}(1-\theta)^{1-x_1}\theta^{x_2}(1-\theta)^{1-x_2} = \underbrace{\theta^{x_1+x_2}(1-\theta)^{2-(x_1+x_2)}}_{\uparrow \ \text{This is } g(x_1+x_2)} \cdot \underset{\uparrow \ \text{This is } h(x_1,x_2).}{1} \end{aligned}$$

(Sometimes h is not just 1; for example, that happens with independent observations from a Poisson distribution.)

Another way relies on the definition of sufficiency and the fact that $X_1 + X_2 \mid \theta \sim \text{Bin}(2, \theta)$:

$$\begin{aligned} &\mathbf{Pr}_\theta(X_1 = x_1 \ \& \ X_2 = x_2 \mid X_1 + X_2 = x_1 + x_2) \\ &= \frac{\mathbf{Pr}_\theta(X_1 = x_1 \ \& \ X_2 = x_2 \ \& \ X_1 + X_2 = x_1 + x_2)}{\mathbf{Pr}_\theta(X_1 + X_2 = x_1 + x_2)} = \frac{\mathbf{Pr}_\theta(X_1 = x_1 \ \& \ X_2 = x_2)}{\mathbf{Pr}_\theta(X_1 + X_2 = x_1 + x_2)} \\ &= \frac{\theta^{x_1+x_2}(1-\theta)^{2-(x_1+x_2)}}{\binom{2}{x_1+x_2}\theta^{x_1+x_2}(1-\theta)^{2-(x_1+x_2)}} = \frac{1}{\binom{2}{x_1+x_2}} \end{aligned}$$

and this does not depend on θ.

(c) Rely on the Rao-Blackwell theorem to find a better unbiased estimator of θ than the one considered above.

Answer: The better estimator is $\mathbf{E}(X_1 \mid X_1 + X_2)$. We have:

$$\mathbf{E}(X_1 \mid X_1 + X_2 = x) = \mathbf{Pr}(X_1 = 1 \mid X_1 + X_2 = x) = \frac{\mathbf{Pr}(X_1 = 1 \ \& \ X_1 + X_2 = x)}{\mathbf{Pr}(X_1 + X_2 = x)}$$

$$= \frac{\mathbf{Pr}(X_1 = 1 \ \& \ X_2 = x - 1)}{\mathbf{Pr}(X_1 + X_2 = x)} = \frac{\mathbf{Pr}(X_1 = 1)\,\mathbf{Pr}(X_2 = x - 1)}{\mathbf{Pr}(X_1 + X_2 = x)} = \frac{\theta \cdot \mathbf{Pr}(X_2 = x - 1)}{\binom{2}{x}\theta^x(1-\theta)^{2-x}}$$

$$= \frac{\theta \cdot \left\{ \begin{array}{ll} 0 & \text{if } x = 0 \\ 1 - \theta & \text{if } x = 1 \\ \theta & \text{if } x = 2 \end{array} \right\}}{\binom{2}{x}\theta^x(1-\theta)^{2-x}} = \left\{ \begin{array}{ll} 0 & \text{if } x = 0 \\ 1/2 & \text{if } x = 1 \\ 1 & \text{if } x = 2 \end{array} \right\} = \left\{ \begin{array}{ll} 0/2 & \text{if } x = 0 \\ 1/2 & \text{if } x = 1 \\ 2/2 & \text{if } x = 2 \end{array} \right\} = \frac{x}{2}.$$

Therefore the estimator that we seek is

$$\mathbf{E}(X_1 \mid X_1 + X_2) = \frac{X_1 + X_2}{2}.$$