Chapter 15 Solutions

Andrew Wu
Wasserman: All of Statistics

August 9, 2025

Problem 15.1. Suppose that Y and Z are both binary and consider data $(Y_1, Z_1), \dots, (Y_n, Z_n)$. Let X_{ij} represent the number of observations for which $Z = i$ and $Y = j$, and let $X_{i\cdot} = \sum_j X_{ij}$, $X_{\cdot j} = \sum_i X_{ij}$, with $n = X_{\cdot\cdot} = \sum_{ij} X_{ij}$. Set $p_{ij} = \mathbb{P}(Z = i, Y = j)$, and let $p_{\cdot j} = \sum_i p_{ij}$ and $p_{i\cdot} = \sum_j p_{ij}$.

Define $\psi = \frac{p_{00}p_{11}}{p_{01}p_{10}}$ and $\gamma = \log \psi$.

Show that the following statements are equivalent:

1. $Y \perp\!\!\!\perp Z$.
2. $\psi = 1$.
3. $\gamma = 0$.
4. For $i, j \in \{0, 1\}$, $p_{ij} = p_{i\cdot}p_{\cdot j}$.

Solution. Note firstly that $p_{\cdot j} = \sum_i p_{ij} = \sum_i \mathbb{P}(Z = i, Y = j) = \mathbb{P}(Y = j)$, and similarly $p_{i\cdot} = \mathbb{P}(Z = i)$.

Now assume statement 1, $Y \perp\!\!\!\perp Z$. Then $\mathbb{P}(Y = y, Z = z) = \mathbb{P}(Y = y)\mathbb{P}(Z = z)$. It follows that

$$\begin{aligned}
\psi &= \frac{p_{00}p_{11}}{p_{01}p_{10}} \\
&= \frac{\mathbb{P}(Z = 0, Y = 0)\mathbb{P}(Z = 1, Y = 1)}{\mathbb{P}(Z = 0, Y = 1)\mathbb{P}(Z = 1, Y = 0)} \\
&= \frac{\mathbb{P}(Z = 0)\mathbb{P}(Y = 0)\mathbb{P}(Z = 1)\mathbb{P}(Y = 1)}{\mathbb{P}(Z = 0)\mathbb{P}(Y = 1)\mathbb{P}(Z = 1)\mathbb{P}(Y = 0)} \\
&= 1
\end{aligned}$$

so we get that statement 1 implies statement 2.

Then, from $\psi = 1$ we immediately obtain $\gamma = \log \psi = 0$, and conversely if $\gamma = 0$ then $\psi = e^0 = 1$, so statements 2 and 3 are equivalent.

Now assume statement 2 (and 3). We will show statement 4. We have

$$\begin{aligned}
p_{00} &= \frac{p_{01}p_{10}}{p_{11}} \\
&= \frac{(p_{0\cdot} - p_{00})(p_{\cdot 0} - p_{00})}{1 - p_{0\cdot} - p_{10}}
\end{aligned}$$

so multiplying through, we get

$$p_{00} - p_{00}p_{0\cdot} - p_{00}(p_{\cdot 0} - p_{00}) = p_{0\cdot}p_{\cdot 0} - p_{00}p_{\cdot 0} - p_{00}p_{0\cdot} + p_{00}^2$$

from which the identity $p_{00} = p_{0\cdot}p_{\cdot 0}$ becomes clear upon some expansion and cancellation. Symmetrically, we obtain $p_{11} = p_{1\cdot}p_{\cdot 1}$.

Then we have

$$\begin{aligned}
p_{01} &= \frac{p_{00}p_{11}}{p_{10}} \\
&= \frac{p_{0\cdot}p_{\cdot 0}p_{1\cdot}p_{\cdot 1}}{p_{1\cdot} - p_{11}} \\
&= \frac{p_{0\cdot}p_{\cdot 0}p_{1\cdot}p_{\cdot 1}}{p_{1\cdot} - p_{1\cdot}p_{\cdot 1}} \\
&= \frac{p_{0\cdot}p_{\cdot 0}p_{\cdot 1}}{1 - p_{\cdot 1}} \\
&= \frac{p_{0\cdot}p_{\cdot 0}p_{\cdot 1}}{p_{\cdot 0}} = p_{0\cdot}p_{\cdot 1}
\end{aligned}$$

as desired. Similarly we get $p_{10} = p_{1\cdot}p_{\cdot 0}$, and thus statements 2 and 3 imply statement 4.

Evidently, statement 4 translates to $\mathbb{P}(Z = i, Y = j) = \mathbb{P}(Z = i)\mathbb{P}(Y = j)$ for all i, j, which gives us $Y \perp\!\!\!\perp Z$. Thus we have established a chain of implications $1 \implies 2 \implies 3 \implies 4 \implies 1$, and so statements $1, 2, 3, 4$ are equivalent. (That is, if any of them are true all of them are true, and if any of them are false than none of the others can be true, and thus must be false.) $\square$

Problem 15.2. Consider testing $H_0 : Y \perp\!\!\!\perp Z$ versus $H_1 : Y \not\perp\!\!\!\perp Z$.

Note that $X \sim \text{Multinomial}(n, p)$ with $p = (p_{00}, p_{01}, p_{10}, p_{11})$ and $X = (X_{00}, X_{01}, X_{10}, X_{11})$ denoting the vector of counts.

Let the likelihood ratio test statistic be T. Show that

$$T = 2 \sum_{i=0}^{1} \sum_{j=0}^{1} X_{ij} \log \left(\frac{X_{ij} X_{\cdot\cdot}}{X_{i\cdot} X_{\cdot j}} \right).$$

Show that $T \xrightarrow{\text{d}} \chi_1^2$.

Solution. The likelihood ratio statistic is

$$T = 2 \log \frac{\mathcal{L}(\widehat{p})}{\mathcal{L}(\widehat{p}_0)}.$$

Here, $\widehat{p}$ is known to be the empirical proportions; that is, $\widehat{p} = (\frac{X_{00}}{n}, \frac{X_{01}}{n}, \frac{X_{10}}{n}, \frac{X_{11}}{n})$, or $\widehat{p} = \frac{X}{n}$. To compute $\widehat{p}_0$, we need the MLE under H_0; that is, the MLE assuming $Y \perp\!\!\!\perp Z$, or equivalently, the odds ratio $\psi = 1$.

Note that the likelihood is

$$\mathcal{L}(p) = \binom{n}{X_{00}, X_{01}, X_{10}, X_{11}} p_{00}^{X_{00}} p_{01}^{X_{01}} p_{10}^{X_{10}} p_{11}^{X_{11}},$$

so dropping constants the log-likelihood is $\ell(p) = X_{00} \log p_{00} + X_{01} \log p_{01} + X_{10} \log p_{10} + X_{11} \log p_{11}$. We want to maximize this under our independence and $p_{00} + p_{01} + p_{10} + p_{11} = 1$ constraints.

Let $u_i = p_{i\cdot}$ and $v_i = p_{\cdot i}$ be the marginals. We will now reparametrize, as such:

$$\begin{aligned}
\ell(p) &= X_{00} \log p_{00} + X_{01} \log p_{01} + X_{10} \log p_{10} + X_{11} \log p_{11} \\
&= X_{00} \log(u_0 v_0) + X_{01} \log(u_0 v_1) + X_{10} \log(u_1 v_0) + X_{11} \log(u_1 v_1) \\
&= (X_{00} + X_{01}) \log u_0 + (X_{10} + X_{11}) \log u_1 + (X_{00} + X_{10}) \log v_0 + (X_{01} + X_{11}) \log v_1
\end{aligned}$$

and we need to maximize this under the constraints that $u_0 + u_1 = 1$ and $v_0 + v_1 = 1$. That means we can simply maximize the first two terms and the last two terms separately.

We have

$$(X_{00} + X_{01}) \log u_0 + (X_{10} + X_{11}) \log u_1 = (X_{00} + X_{01}) \log u_0 + (X_{10} + X_{11}) \log(1 - u_0)$$

and taking the derivative and setting it equal to 0, we get

$$\frac{X_{10} + X_{11}}{1 - u_0} = \frac{X_{00} + X_{01}}{u_0}$$

and simplifying,

$$u_0 = \frac{X_{00} + X_{01}}{n}.$$

Similarly, $u_1 = \frac{X_{10}+X_{11}}{n}$, $v_0 = \frac{X_{00}+X_{10}}{n}$, and $v_1 = \frac{X_{01}+X_{11}}{n}$. From these, we recover the estimates

$$\widehat{p}_{ij} = \frac{X_{i\cdot} X_{\cdot j}}{n^2}.$$

Thus, substituting everything back in for our formula for T, we have

$$\begin{aligned} 2\log\frac{\mathcal{L}(\widehat{p})}{\mathcal{L}(\widehat{p}_0)} &= 2\log\left(\frac{(\frac{X_{00}}{n})^{X_{00}}(\frac{X_{01}}{n})^{X_{01}}(\frac{X_{10}}{n})^{X_{10}}(\frac{X_{11}}{n})^{X_{11}}}{(\frac{X_{0\cdot}X_{\cdot 0}}{n^2})^{X_{00}}(\frac{X_{0\cdot}X_{\cdot 1}}{n^2})^{X_{01}}(\frac{X_{1\cdot}X_{\cdot 0}}{n^2})^{X_{10}}(\frac{X_{1\cdot}X_{\cdot 1}}{n^2})^{X_{11}}}\right) \\ &= 2\sum_{i=0}^{1}\sum_{j=0}^{1}\log\left[\left(\frac{X_{ij}n}{X_{i\cdot}X_{\cdot j}}\right)^{X_{ij}}\right] = 2\sum_{i=0}^{1}\sum_{j=0}^{1} X_{ij}\log\left(\frac{X_{ij}X_{\cdot\cdot}}{X_{i\cdot}X_{\cdot j}}\right). \end{aligned}$$

Finally, we know that $T \xrightarrow{\text{d}} \chi^2_1$, as there is but one additional constraint imposed between H_1 and H_0 (the constraint on $p_{00}p_{11} = p_{01}p_{10}$.) We are done. □

Problem 15.3. Prove that the MLEs of ψ and γ (as they are defined in problem 15.1) are

$$\widehat{\psi} = \frac{X_{00}X_{11}}{X_{01}X_{10}}, \ \ \widehat{\gamma} = \log\widehat{\psi}.$$

Prove that the asymptotic standard errors, computed using the delta method, are

$$\begin{aligned} \widehat{\text{se}}(\widehat{\gamma}) &= \sqrt{\frac{1}{X_{00}} + \frac{1}{X_{01}} + \frac{1}{X_{10}} + \frac{1}{X_{11}}} \\ \widehat{\text{se}}(\widehat{\psi}) &= \widehat{\psi}\cdot\widehat{\text{se}}(\widehat{\gamma}). \end{aligned}$$

Solution. We know that $\psi = \frac{p_{00}p_{11}}{p_{01}p_{10}}$, so $\widehat{\psi} = \frac{\widehat{p}_{00}\widehat{p}_{11}}{\widehat{p}_{01}\widehat{p}_{10}}$. We also know that $\widehat{p}_{ij} = \frac{X_{ij}}{n}$. It follows by equivariance properties that $\widehat{\psi} = \frac{X_{00}X_{11}}{X_{01}X_{10}}, \widehat{\gamma} = \log\widehat{\psi}$.

Now we compute the standard errors. First, suppose that $\gamma = g(p) = \log p_{00} + \log p_{11} - \log p_{01} - \log p_{10}$. Then

$$\nabla g = \begin{pmatrix} \frac{1}{p_{00}} \\ -\frac{1}{p_{01}} \\ -\frac{1}{p_{10}} \\ \frac{1}{p_{11}} \end{pmatrix}.$$

Moreover, we know that $\mathbb{V}(X_{ij}) = np_{ij}(1-p_{ij})$, and

$$\begin{aligned} \text{Cov}(X_{ij}, X_{kl}) &= \mathbb{E}(X_{ij}X_{kl}) - \mathbb{E}(X_{ij})\mathbb{E}(X_{kl}) \\ &= \mathbb{E}(X_{ij}X_{kl}) - n^2 p_{ij}p_{kl}. \end{aligned}$$

Let us now compute $\mathbb{E}(X_{ij}X_{kl})$. Let W be the result of a arbitrary trial, and let I_{ij} be the indicator function that returns 1 if W landed in the ij case. That is, $I_{10}(W) = 1$ if W contributes to X_{10}, and 0 otherwise.

Then, noting that $I_{ij}(W_u)I_{kl}(W_u) = 0$, we have

$$\begin{aligned} \mathbb{E}(X_{ij}X_{kl}) &= \mathbb{E}\left[\left(\sum_{u=1}^{n} I_{ij}(W_u)\right)\left(\sum_{v=1}^{n} I_{kl}(W_v)\right)\right] \\ &= \mathbb{E}\left[\sum_{u\neq v} I_{ij}(W_u)I_{kl}(W_v)\right] \\ &= n(n-1)\mathbb{E}[I_{ij}(W_u)I_{kl}(W_v)] = n(n-1)p_{ij}p_{kl} \end{aligned}$$

so $\mathrm{Cov}(X_{ij}, X_{kl}) = -np_{ij}p_{kl}$.

Note that this gives us the variance-covariance matrix for the counts X_{ij}, so to get the variance-covariance matrix for the estimators, we simply divide everything by n^2, as the estimators are the counts divided by n. Then

$$\mathbb{V}(\widehat{p}) = \frac{1}{n}\begin{bmatrix} p_{00}(1-p_{00}) & -p_{00}p_{01} & -p_{00}p_{10} & -p_{00}p_{11} \\ -p_{01}p_{00} & p_{01}(1-p_{01}) & -p_{01}p_{10} & -p_{01}p_{11} \\ -p_{10}p_{00} & -p_{10}p_{01} & p_{10}(1-p_{10}) & -p_{10}p_{11} \\ -p_{11}p_{00} & -p_{11}p_{01} & -p_{11}p_{10} & p_{11}(1-p_{11}) \end{bmatrix}.$$

Now we must compute $\mathbb{V}(\widehat{\gamma}) = \nabla(g)(\widehat{p})^T\mathbb{V}(\widehat{p})\nabla(g)(\widehat{p})$. Let $M = \frac{1}{n}\mathbb{V}(\widehat{p})$, let $\mathrm{col}_i(M)$ denote the ith column of M, let $\nabla g = \nabla g(\widehat{p})$, and let ∇g_i denote the ith element of ∇g. Then,

$$\begin{aligned} \mathbb{V}(\widehat{\gamma}) &= \sum_{i=1}^{4}(\nabla g^T \cdot \mathrm{col}_i(M)) \cdot \nabla g_i \\ &= \sum_{i=1}^{4}\left(\sum_{j=1}^{4}\nabla g_j M_{ji}\nabla g_i\right) \\ &= \sum_{(i,j)}\nabla g_i \nabla g_j M_{ij}. \end{aligned}$$

In the summation above, when $i = j$, we get the sum $\frac{1}{n}\sum_{(i,j)}\frac{1-p_{ij}}{p_{ij}}$. When $i \neq j$, we have two cases: in the case that $i = 2, j = 3$ and $i = 1, j = 4$, we get $M_{ij}\nabla g_i \nabla g_j = -\frac{1}{n}$, and otherwise we get $M_{ij}\nabla g_i\nabla g_j = \frac{1}{n}$. So we have

$$\begin{aligned} \mathbb{V}(\widehat{\gamma}) &= \frac{1}{n}\left(\frac{1-p_{00}}{p_{00}} + \frac{1-p_{01}}{p_{01}} + \frac{1-p_{10}}{p_{10}} + \frac{1-p_{11}}{p_{11}} + 8 - 4\right) \\ &= \frac{1}{n}\left(\frac{1}{p_{00}} + \cdots + \frac{1}{p_{11}}\right) \end{aligned}$$

and using $\widehat{p}_{ij} = \frac{X_{ij}}{n}$, we obtain

$$\widehat{\mathrm{se}}(\widehat{\gamma}) = \sqrt{\frac{1}{X_{00}} + \frac{1}{X_{01}} + \frac{1}{X_{10}} + \frac{1}{X_{11}}}$$

as desired.

Next, we use the fact that $\psi = e^{\gamma}$, giving us

$$\mathrm{se}(\widehat{\psi}) = |e^{\widehat{\gamma}}|\mathrm{se}(\widehat{\gamma})$$

so we just need to show that $\widehat{\psi} = |e^{\widehat{\gamma}}|$, but we already showed that $\widehat{\gamma} = \log\widehat{\psi}$, so we are done. $\square$

Problem 15.4. Consider the following data on death sentencing and race:

	Death Sentence	No Death Sentence
Black Victim	14	641
White Victim	62	594

Analyze the data using the tools from this chapter. Interpret the results. Explain why, based only on this information, you can't make causal conclusions.

Solution. Let $Y = 0$ if a victim was given the death sentence, and $Y = 1$ if not. Let $Z = 0$ if the victim was black, and $Z = 1$ if the victim was white. Let X_{ij} be the number of observations for which $Z = i$ and $Y = j$. We'll compute the likelihood ratio test statistic, testing $H_0 : Y \perp\!\!\!\perp Z$ versus $H_1 : Y \not\perp\!\!\!\perp Z$ with $\alpha = 0.05$.

We have

$$
\begin{aligned}
T &= 2\sum_{i=0}^{1}\sum_{j=0}^{1} X_{ij}\log\left(\frac{X_{ij}X_{\cdot\cdot}}{X_{i\cdot}X_{\cdot j}}\right) \\
&= 2\left(14\log\left(\frac{14\cdot 1311}{655\cdot 76}\right) + 62\log\left(\frac{62\cdot 1311}{76\cdot 656}\right) + 641\log\left(\frac{641\cdot 1311}{655\cdot 1235}\right) + 594\log\left(\frac{594\cdot 1311}{656\cdot 1235}\right)\right) \\
&\approx 34.53.
\end{aligned}
$$

Then $\chi^2_{1,\alpha} \approx 3.84$, so $T > \chi^2_{1,\alpha}$ and we reject the null hypothesis. We conclude that Y and Z are associated; that is, having a death sentence is associated with race, but this tool does not allow us to conclude anything about causation. $\square$