

# Chapter 8 Sols

- ① Go to the Jupyter notebook for Lecture 17.  
 On cell 10, substitute `np.linspace(1, 11, 200)` with  
`np.linspace(1, 11, 11)`.

Results for  $\hat{P}_{\text{Die} \mid \text{Group}} = [0.03, 0.06, 0.12,$   
 $0.22, 0.38, 0.57, 0.74, 0.86, 0.93, 0.97, 0.98]$ .  
 If we want  $E(\text{Die} \mid \text{Group})$ , just multiply times 10,  
 $\pi_{11}$ . (observed counts in blue)

If we want  $E(\text{Die} | \text{Group})$ , <sup>just write</sup>  $\underline{\text{observed counts}}$  in blue  
 +6 # of mice in each group. Thus,  $E(\text{Die} | \text{Group}) = [0.3, 0.6, 1.2, 2.2, 3.8, 5.7, 7.4,$   
 $8.6, 9.3, 9.7, 9.8]$

$$\textcircled{2} f(y|\pi) = \binom{n}{y} \pi^y (1-\pi)^{n-y}$$

Let's get  $L_y(\pi)$  in order to gain intuition about what  
 $\lambda$  should be.

$$L_y(\pi) = y \cdot \log(\pi) + (n-y) \log(1-\pi)$$

λ show:

$$\begin{aligned}
 L_y(\pi) &= y \cdot \log(\pi) + (n-y) \log(1-\pi) \\
 &= y \cdot (\log(\pi) - \log(1-\pi)) + n \log(1-\pi) \\
 &= y \cdot \log\left(\frac{\pi}{1-\pi}\right) + n \log(1-\pi)
 \end{aligned}$$

Let  $\lambda = \log\left(\frac{\pi}{1-\pi}\right)$  and

$$\begin{aligned}
 e^\lambda &= \frac{\pi}{1-\pi} = e^\lambda - e^\lambda \cdot \pi = \pi \Rightarrow \pi(1+e^\lambda) = e^\lambda \Rightarrow \pi = \frac{e^\lambda}{1+e^\lambda} \\
 \Rightarrow n \cdot \log(1-\pi) &= n \cdot \log\left(1 - \frac{e^\lambda}{1+e^\lambda}\right) = n \cdot \log\left(\frac{1}{1+e^\lambda}\right) \\
 &= n \cdot [\log(1) - \log(1+e^\lambda)]
 \end{aligned}$$

1 -  $\pi$ 1 +  $e^{\lambda}$ 

$$\Rightarrow n \cdot \log(1 - \pi) = n \cdot \log\left(1 - \frac{e^\lambda}{1 + e^\lambda}\right) = n \cdot \log\left(\frac{1}{1 + e^\lambda}\right)$$

$$= n \cdot [\log(1) - \log(1 + e^\lambda)]$$

$$= -n \cdot \log(1 + e^\lambda)$$

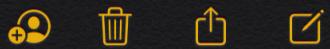
Thus,  $\ell_y(\lambda) = y \cdot \lambda - n \cdot \log(1 + e^\lambda)$ . Finally, taking the exponential

$$f(y|\lambda) = e^{y\lambda - n\ell(\lambda)} \binom{n}{y}, \text{ where } \ell(x) = \log(1 + e^x)$$



— . . . —





③ See Jupyter Notebook.

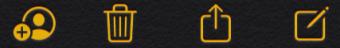
④ Departing from the general definition  
of Deviance (formula 0.31) ...

$$D(f_1, f_2) = 2 \int_{-\infty}^{\infty} f_1(y) \cdot \log \frac{f_1(y)}{f_2(y)} \cdot dy$$

and letting  $y \sim \text{Poi}(\mu)$

$$\Rightarrow D(f_1, f_2) = -2 \left[ \sum_{y=0}^{\infty} \mu^y e^{-\mu} \cdot \log \left( \frac{\mu^y e^{-\mu}}{\dots^y e^{-\mu}} \right) \right]$$





$$D(f_1, f_2) = 2 \int_Y f_1(y) \cdot \log \frac{f_1(y)}{f_2(y)} \cdot dy$$

and letting  $y \sim \text{Poi}(\mu)$

$$\Rightarrow D(f_1, f_2) = 2 \cdot \left[ \sum_{y=0}^{\infty} \frac{\mu_1^y \cdot e^{-\mu_1}}{y!} \cdot \log \left( \frac{\mu_1^y \cdot e^{-\mu_1}}{\mu_2^y \cdot e^{-\mu_2}} \right) \right]$$

$$= 2 \cdot \left[ \sum_{y=0}^{\infty} \frac{\mu_1^y \cdot e^{-\mu_1}}{y!} \cdot \log \left( \frac{\mu_1}{\mu_2} \right)^y + \log \left( \frac{e^{-\mu_1}}{e^{-\mu_2}} \right) \right]$$

$$= 2 \cdot \left[ \sum_{y=0}^{\infty} \frac{\mu_1^y \cdot e^{-\mu_1}}{y!} \cdot \left[ y \cdot \log \left( \frac{\mu_1}{\mu_2} \right) - (\mu_1 - \mu_2) \right] \right]$$

$$= 2 \cdot \left[ \mu_1^y \cdot e^{-\mu_1} \Big|_{y=0} (\mu_1) - \sum_{y=0}^{\infty} \mu_1^y \cdot e^{-\mu_1} (\mu_1 - \mu_2) \right]$$





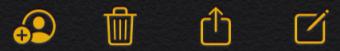
$$= 2 \cdot \left[ \sum_{y=0}^{\infty} \frac{\mu_1^y \cdot e^{-\mu_1}}{y!} \cdot \left[ y \cdot \log\left(\frac{\mu_1}{\mu_2}\right) - (\mu_1 - \mu_2) \right] \right]$$

$$= 2 \cdot \underbrace{\left[ \sum_{y=0}^{\infty} \frac{\mu_1^y \cdot e^{-\mu_1}}{y!} \cdot y \cdot \log\left(\frac{\mu_1}{\mu_2}\right) \right]}_{\text{Sum of } f \text{ Pois}(\mu_1) \text{ pmf.}} - \underbrace{\sum_{y=0}^{\infty} \frac{\mu_1^y \cdot e^{-\mu_1}}{y!} (\mu_1 - \mu_2)}$$

Expected value of  $Pois(\mu_1) = \mu_1$

$$= 2 \cdot \left[ \mu_1 \cdot \log\left(\frac{\mu_1}{\mu_2}\right) - (\mu_1 - \mu_2) \right]$$

$$= 2 \mu_1 \cdot \left[ \log\left(\frac{\mu_1}{\mu_2}\right) - \left( 1 - \frac{\mu_2}{\mu_1} \right) \right]$$



- expect value - - - - , , ,

$$= 2 \cdot \left[ w_1 \cdot \log\left(\frac{w_1}{w_2}\right) - (w_1 - w_2) \right]$$

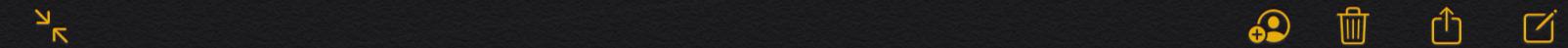
$$= 2 w_1 \cdot \left[ \log\left(\frac{w_1}{w_2}\right) - \left(1 - \frac{w_2}{w_1}\right) \right]$$

$$= 2 w_1 \cdot \left[ \left( \frac{w_2}{w_1} - 1 \right) - \log\left(\frac{w_2}{w_1}\right) \right] .$$

⑤ (8.28) states  $x^T [y - w(\hat{\alpha})] = 0$

$$\Rightarrow x^T y - x^T w(\hat{\alpha}) = 0 \Rightarrow z - E(\hat{z}) = 0$$





⑤ (8.28) states  $x^T [y - \mu(\hat{\alpha})] = 0$

$$\Rightarrow x^T y - x^T \mu(\hat{\alpha}) = 0 \Rightarrow z - E(\hat{z}) = 0$$

This, the optimality condition 8.28 states that  
The likelihood maximizing choice of  $\hat{\alpha}$  is that which  
produces estimate  $\hat{z} = x^T \mu(\hat{\alpha})$  s.t.  $E(\hat{z}) = z$ .  
(i.e.  $\hat{z}$  is unbiased).

⑥ See Jupyter Notebook.





6

See Jupyter Notebook.

7

$$S^2_k = \sum_{y \in k} (y_i - m_k)^2$$

$$S^2_k = \sum_{y \in k, l} (y_i - m_k)^2 + \sum_{y \in k, r} (y_i - m_k)^2$$

$$= \sum_{y \in k, l} ((y_i - m_{k,l}) + (m_{k,l} - m_k))^2 + \sum_{y \in k, r} ((y_i - m_{k,r}) + (m_{k,r} - m_k))^2$$

$$= \underbrace{\sum_{y \in k, l} (y_i - m_{k,l})^2}_{S^2_{k,l}} + \underbrace{\left[ 2 \cdot \sum_{y \in k, l} (y_i - m_{k,l})(m_{k,l} - m_k) \right]}_{\text{cross terms}} + \sum_{y \in k, l} (m_{k,l} - m_k)^2$$

$$+ \underbrace{\sum_{y \in k, r} (y_i - m_{k,r})^2}_{\text{cross terms}} + \underbrace{\left[ 2 \cdot \sum_{y \in k, r} (y_i - m_{k,r})(m_{k,r} - m_k) \right]}_{\text{cross terms}} + \sum_{y \in k, r} (m_{k,r} - m_k)^2$$

5 ②

① 📷 ⌂

$$\begin{aligned}
 &= \sum_{y \in k/l} ((y_i - m_{k/l}) + (m_{k/l} - m_k)) + \sum_{y \in k/r} ((y_i - m_{k/r}) + (m_{k/r} - m_k)) \\
 &= \sum_{y \in k/l} (y_i - m_{k/l})^2 + \left[ 2 \cdot \sum_{y \in k/l} (y_i - m_{k/l})(m_{k/l} - m_k) \right] + \sum_{y \in k/l} (m_{k/l} - m_k)^2 \\
 &\quad + \sum_{y \in k/r} (y_i - m_{k/r})^2 + \left[ 2 \cdot \sum_{y \in k/r} (y_i - m_{k/r})(m_{k/r} - m_k) \right] + \sum_{y \in k/r} (m_{k/r} - m_k)^2
 \end{aligned}$$

Note that  $(m_{k/l} - m_k)$  are constants in these two sums.  
 Since we know  $\sum (y_i - \bar{y}) = 0$

$$\Rightarrow (m_{k/l} - m_k) \cdot \sum_{y \in k/l} (y_i - m_{k/l}) = 0$$

$$(m_{k/r} - m_k) \cdot \sum_{y \in k/r} (y_i - m_{k/r}) = 0$$

$$\Rightarrow S^2_K = S^2_{k/l} + S^2_{k/r} + \sum_{y \in k/l} (m_{k/l} - m_k)^2 + \sum_{y \in k/r} (m_{k/r} - m_k)^2$$

$$\approx \dots + N_{k/l} (m_{k/l} - m_k)^2 + N_{k/r} (m_{k/r} - m_k)^2$$



$$\begin{aligned}
 & \Rightarrow (m_{kr} - m_k) \cdot \sum_{x \in K \setminus r} (\gamma_i - m_{xr}) = 0 \\
 \Rightarrow S^2_K &= S^2_{k,l} + S^2_{k,r} + \sum_{y \in K \setminus l} (m_{yl} - m_k)^2 + \sum_{y \in K \setminus r} (m_{yr} - m_k)^2 \\
 &= S^2_{k,l} + S^2_{k,r} + N_{kl} (m_{kl} - m_k)^2 + N_{kr} (m_{kr} - m_k)^2 \\
 &= S^2_{k,l} + S^2_{k,r} + N_{kl} (m_{kl} - m_k)^2 + N_{kr} (m_k - m_{kr})^2 \\
 &= S^2_{k,l} + S^2_{k,r} + 2N_{kl} \cdot m_k \cdot m_{kl} + N_{kl} \cdot m_k^2 \\
 &\quad + N_{kr} \cdot m_{kl} - 2N_{kl} \cdot m_k \cdot m_{kr} + N_{kr} \cdot m_{kr}^2 \\
 &\quad + N_{kr} \cdot m_k^2 - 2N_{kr} \cdot m_k \cdot m_{kr} + N_{kr} \cdot m_{kr}^2 \\
 &= S^2_{k,l} + S^2_{k,r} \\
 &\quad + N_{kl} \cdot m_{kl} + N_{kr} \cdot m_{kr} - 2m_k (N_{kl} \cdot m_{kl} + N_{kr} \cdot m_{kr})
 \end{aligned}$$





T 1W m

$$= S_{K,l}^2 + S_{K,r}^2 \\ + N_{Ke} \cdot m_{Ke}^2 + N_{Kr} \cdot m_{Kr}^2 - 2 m_K (N_{Ke} \cdot m_{Ke} + N_{Kr} \cdot m_{Kr}) \\ + (N_{Ke} + N_{Kr}) \cdot m_K^2$$

$$= S_{K,l}^2 + S_{K,r}^2 \\ + N_{Ke} \cdot m_{Ke}^2 + N_{Kr} \cdot m_{Kr}^2 - 2 \frac{(N_{Ke} \cdot m_{Ke} + N_{Kr} \cdot m_{Kr})^2}{N_K} \\ + \cancel{\frac{(N_{Ke} + N_{Kr}) \cdot (N_{Ke} \cdot m_{Ke} + N_{Kr} \cdot m_{Kr})^2}{N_K^2}}$$

$$= S_{K,l}^2 + S_{K,r}^2 \\ + N_{Ke} \cdot m_{Ke}^2 + N_{Kr} \cdot m_{Kr}^2 - \frac{(N_{Ke} \cdot m_{Ke} + N_{Kr} \cdot m_{Kr})^2}{N_K}$$





$$= S_{\gamma, l}^2 + S_{\gamma, r}^2$$

$$+ \frac{N_K (N_{K, l} \cdot m_{Kl}^2 + N_{K, r} \cdot m_{Kr}^2) - (N_{Kl} \cdot m_{Kl} + N_{Kr} \cdot m_{Kr})^2}{N_K}$$

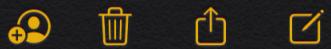
$$= S_{\gamma, l}^2 + S_{\gamma, r}^2$$

$$+ \frac{(N_{K, l} + N_{K, r})(N_{K, l} \cdot m_{Kl}^2 + N_{K, r} \cdot m_{Kr}^2) - (N_{Kl} \cdot m_{Kl} + N_{Kr} \cdot m_{Kr})^2}{N_K}$$

$$= S_{\gamma, l}^2 + S_{\gamma, r}^2 +$$

$$\frac{1}{N_K} \cdot \left[ \cancel{N_{\gamma, l} m_{\gamma, l}^2} + N_{\gamma, l} \cdot N_{\gamma, r} m_{\gamma, r}^2 + \cancel{N_{\gamma, r} \cdot N_{\gamma, l} m_{\gamma, l}^2} + \cancel{N_{\gamma, r} m_{\gamma, r}^2} \right] \\ - \cancel{N_{\gamma, l}^2 m_{\gamma, l}^2} = 2 N_{\gamma, l} \cdot N_{\gamma, r} m_{\gamma, l} \cdot m_{\gamma, r} - \cancel{N_{\gamma, r}^2 m_{\gamma, r}^2}$$



 $\approx$ 

$$\begin{aligned}
 &= S_{K,l}^2 + S_{K,r}^2 + \\
 &\frac{1}{N_K} \cdot \left[ N_{K,l} m_{K,l}^2 + N_{K,r} m_{K,r}^2 + N_{K,r} \cdot N_{K,l} \cdot m_{K,l}^2 + N_{K,r} m_{K,r}^2 \right. \\
 &\quad \left. - N_{K,l} m_{K,l}^2 = 2N_{K,r} \cdot N_{K,l} m_{K,l} \cdot m_{K,r} - N_{K,r}^2 m_{K,r}^2 \right] \\
 &= S_{K,l}^2 + S_{K,r}^2 + \frac{N_{K,l} \cdot N_{K,r}}{N_K} \cdot \left( m_{K,l}^2 - 2m_{K,l} m_{K,r} + m_{K,r}^2 \right) \\
 \Rightarrow S_K^2 &= S_{K,l}^2 + S_{K,r}^2 + \frac{N_{K,l} \cdot N_{K,r}}{N_K} \cdot \left( m_{K,l}^2 - m_{K,r}^2 \right)^2
 \end{aligned}$$