1.13 The counts follow a multinomial dist. given \( \vec{\nu} = (\nu_1, \nu_2, \nu_3, \nu_4, \nu_5) \)
\[ n = \sum_{i=1}^{5} n_i = 28. \quad \text{That is,} \]
\[ f(n_1, \ldots, n_5 | \vec{\nu}) = \frac{n!}{n_1! \ldots n_5!} \nu_1^{n_1} \ldots \nu_5^{n_5} \]

The prior of the proportion follows a Dirichlet dist. with \( \xi_i = k(5) \)
\[ p(\vec{\nu}) = \frac{\Gamma(k)}{\Gamma(n+1)} \frac{\vec{\nu}^{k-1}}{\nu^1 \nu_2^{q_2} \ldots \nu_5^{q_5}} \]
\[ \Rightarrow \]

posterior: \[ f(n_1, \ldots, n_5 | \vec{\nu}) \propto f(n_1, \ldots, n_5 | \vec{\nu}) \cdot p(\vec{\nu}) \]
\[ \propto \frac{n!}{n_1! \ldots n_5!} \nu_1^{n_1} \ldots \nu_5^{n_5} \]
\[ \Rightarrow \quad L(\vec{\nu}; n_1, \ldots, n_5) = \frac{n_1 + \xi_1}{n + k} \quad i = 1, \ldots, 5 \]

\[ \Rightarrow \quad \text{for } k = 1 \quad \vec{\nu} = (0.007, 0.283, 0.352, 0.317, 0.015) \]
\[ \text{for } k = 2.5 \quad \vec{\nu} = (0.016, 0.279, 0.344, 0.311, 0.089) \quad \hat{\vec{\nu}} \]
\[ \text{for } k = 5 \quad \vec{\nu} = (0.030, 0.273, 0.333, 0.303, 0.061) \]

\[ \hat{\nu}_{\text{MLE}} = \arg \max_{\vec{\nu}} f(n_1, \ldots, n_5 | \vec{\nu}) \quad \Rightarrow \quad \hat{\nu}_{\text{MLE}} = \frac{n_i}{n} \quad \hat{\vec{\nu}} \]

So for "much more" category all MLE's are \( \hat{\vec{\nu}}_{\text{MLE}} = \vec{0} \)

2.2 (a) \( X \) given \( Y \)

b) we want \( P(Y = w | X = w) \)

d) According to the Bayes' Thm. \( P(Y = w | X = w) = \frac{P(X = w | Y = w) \cdot P(Y = w)}{P(X = w | Y = w) \cdot P(Y = w) + P(X = w | Y = B) \cdot P(Y = B)} \)

Time we know \( P(X = w | Y = w) \) and \( P(X = w | Y = B) \)

We only need \( P(Y = w) \) or \( P(Y = B) \)

c) From 2.1 we have \( P(X = F, Y = F) = 10.5\% \), \( P(X = F, Y = M) = 9.5\% \), which are very similar. Then we have.

d) \( P(X = w | Y = w) = 85\% \) and \( P(X = w | Y = B) = 8% \) which are very different.

So the rate of victim and rate of offender have strong association

(Far odds version, \( \theta_{sex} \approx 1.1 \), \( \theta_{race} \approx 6.5 \) which \( \theta_{race} \) is much larger.)
2.4 a) i) \( P(fatal | \text{no}) = \frac{1035}{1035 + 55625} = 0.0191 \)
ii) \( P(fatal | \text{Yes}) = \frac{705}{705 + 441289} = 0.016 \)
b) i) \( P(\text{Use} | \text{fatal}) = \frac{705}{1035 + 705} = 0.393 \)
ii) \( P(\text{Use} | \text{nonfatal}) = \frac{441289}{441289 + 55625} = 0.888 \)

c) \( \tilde{p}_1 = P(fatal | \text{no}) = 0.0191 \quad \tilde{p}_2 = P(fatal | \text{Yes}) = 0.016 \)

\( \Rightarrow \) The difference of proportions of fatal between use seat-belt or not is

\( \tilde{p}_1 - \tilde{p}_2 = 0.0175 \)

The relative risk is \( \frac{\tilde{p}_1}{\tilde{p}_2} = 11.9375 \)

The odds ratio is \( \frac{\tilde{p}_1 \cdot (1 - \tilde{p}_2)}{\tilde{p}_2 \cdot (1 - \tilde{p}_1)} = 12.243 \).

Interpretation:

b) The difference of proportions shows the difference of fatality given wearing a seat belt or not. Here the probability of fatality given not wearing a seat belt is much higher than that wearing a seat belt.

c) The relative risk here means the probability of fatality given not wearing seat belt is about 12 times bigger than that wearing a seat belt.

d) The odds ratio means the odds of fatal for not wearing a seat belt is 12.243 times bigger than that for wearing a seat belt.

The odds ratio is \( \frac{\tilde{p}_1 \cdot (1 - \tilde{p}_2)}{\tilde{p}_2 \cdot (1 - \tilde{p}_1)} = \frac{\tilde{p}_1}{\tilde{p}_2} \cdot \frac{1 - \tilde{p}_2}{1 - \tilde{p}_1} = \text{relative risk} \cdot \frac{1 - \tilde{p}_2}{1 - \tilde{p}_1} \)

Since here \( \frac{1 - \tilde{p}_2}{1 - \tilde{p}_1} \approx 1 \),

the relative risk and odds ratio approximately equal.
2.10 a) The correct interpretation is: The odds of survival for females is 11.4 times the odds for males (not probability).

Like the results in 2.4, let \( T_1 = P(\text{Survival | F}) \) and \( T_2 = P(\text{Survival | M}) \).

The odds ratio \( \theta = \frac{T_1(1-T_2)}{T_2(1-T_1)} = \frac{T_1}{T_2} \cdot \frac{1-T_2}{1-T_1} = \) probability ratio \( \times \frac{1}{\text{odds}} \).

If \( T_1, T_2 \) are very small, \( \frac{1}{\text{odds}} \approx 1 \). The interpretation would be approximately correct.

b) \( \frac{T_1}{T_2} = 2.9 \Rightarrow T_1 = 0.74 \)

\[ \frac{T_1}{T_2} = \frac{2.9}{11.4} = 0.25 \Rightarrow T_2 = 0.2 \]

Thus, the survival proportion for female is 0.74, for male is 0.2.

2.15
- Bag 10B = \( \frac{512 \times 19}{253 \times 89} \approx 0.349 \)
- Bag 11B = \( \frac{353 \times 8}{289 \times 17} \approx 0.805 \)
- Bag 10C = \( \frac{120 \times 321}{205 \times 228} \approx 1.133 \)
- Bag 11D = \( \frac{138 \times 244}{317 \times 137} \approx 0.921 \)
- Bag 10E = \( \frac{53 \times 299}{138 \times 94} \approx 1.222 \)
- Bag 11F = \( \frac{22 \times 317}{351 \times 24} \approx 0.828 \)

The marginal odds ratio: \( \theta_{10} = \frac{1198 \times 1278}{1493 \times 557} \approx 1.84 \)

Interpretation:

The conditional odds ratio means the ratio of the admitted odds for two genders in department \( i \) (i = A, B, .. F) eg. For department A, the odds of admitted of male is 0.3492 times of that of female.

The marginal odds ratio means that for all departments, the odds of admitted of male is 1.841 times of that of female.

The reason they are very different is that the department variable \( D \) has very big effect on admission. And the ratio of gender for different dept. varies a lot.