STA 6384, Report 4.3

Carson Slater *Baylor University*

Problem: Table 1 displays cross-sectional data from a study involving 400 patients with malignant melanoma. Analyze this data using R.

Table 1: Malignant melanoma data from Roberts et al. (1980).

	Tumor Site			
Tumor Type	Head & Neck	Trunk	Extremities	Total
Hutchinson's melanotic freckle	22	2	10	34
Sperficial spreading melanoma	16	54	115	185
Nodular	19	33	73	125
Indeterminant	11	17	28	56
Total	68	106	226	400

The basic idea is to model the log of the expected count in each cell as a function of tumor type, tumor site, and their interaction. The coefficients for the interaction terms will be the **log-odds ratios**. By exponentiating these coefficients and their confidence intervals, we get the odds ratios and their corresponding CIs.

To determine if there's a statistically significant relationship between tumor type and site, we perform a Chi-squared test of independence.

```
# Perform the Pearson's Chi-squared test
chi_test <- chisq.test(melanoma_matrix)
```

```
    Pearson's Chi-squared test

data:  melanoma_matrix
X-squared = 65.813, df = 6, p-value = 2.943e-12
```

We fit a Poisson regression model. The `TumorType * TumorSite` formula tells R to include main effects for both variables and their interaction term.

```
# Fit the Poisson GLM with an interaction term
poisson_model <- glm(
  Count ~ TumorType * TumorSite,
  data = melanoma_df,
  family = poisson
)
```

```
Degrees of Freedom: 11 Total (i.e. Null);  0 Residual
Null Deviance:      295.2
Residual Deviance: 1.021e-14    AIC: 83.11
```

We extract the model coefficients (the log-odds ratios), calculate their confidence intervals, and then exponentiate everything to get the odds ratios (OR) and their 95% CIs. We are only interested in the interaction terms.

	OR	2.5 %	97.5 %
TumorTypeSuperficial:TumorSiteTrunk	37.12	7.87	175.14
TumorTypeNodular:TumorSiteTrunk	19.11	4.04	90.35
TumorTypeIndeterminant:TumorSiteTrunk	17.00	3.32	87.13
TumorTypeSuperficial:TumorSiteExtremities	15.81	6.35	39.37
TumorTypeNodular:TumorSiteExtremities	8.45	3.43	20.83
TumorTypeIndeterminant:TumorSiteExtremities	5.60	2.01	15.57

Because none of the 95% confidence intervals contain 1.0, all of these interaction effects are statistically significant. This confirms the finding from the Chi-squared test that there is a very strong association between melanoma type and its location.

- Interpreting the Odds Ratios (OR): Each OR compares the odds of a tumor being on a specific site (vs. the Head & Neck) for a given tumor type (vs. Hutchinson's).
- For Superficial:Trunk (OR = 37.12), the odds of a tumor being on the Trunk vs. the Head & Neck are about 37 times higher for Superficial melanoma than for Hutchinson's melanoma.
- For Superficial:Extremities (OR = 15.81), the odds of a tumor being on the Extremities vs. the Head & Neck are nearly 16 times higher for Superficial melanoma than for Hutchinson's melanoma.
- Confidence Intervals: While the odds ratios are large, some of the confidence intervals are quite wide (e.g., 7.87 to 175.14 for Superficial:Trunk). This indicates that while the association is very strong, there is some uncertainty about its precise magnitude.

Overall Conclusion

The data provides very strong evidence that Hutchinson's melanoma has a distinct preference for the Head & Neck. In contrast, all other types of melanoma are significantly more likely to be found on the Trunk or Extremities. Superficial spreading melanoma shows this effect most dramatically.