Outline of Solution to Midterm

Problem 1. In a half a page, discuss the advantages of a designed experiment over an observational study.

Problem 2. Recall from toxicological experiments that an important quantity of interest is the median lethal dose, or LD50. For the logistic model, the probability of death is related to dose through,

$$\log(\pi_i/(1-\pi_i)) = \alpha + \beta \log d_i,$$

where π_i is the probability of death and d_i is the dose used in the i-th group; $i = 1, \ldots, k$. Let $(\hat{\alpha}, \hat{\beta})$ be the MLE of (α, β) based on a binomial model in each group. (You do *not* have to know the formula for $\hat{\alpha}$ and $\hat{\beta}$.)

(a) Derive a formula for LD50 in terms of α and β.

Solution. By defination, for $d = \text{LD50}$, $\pi = 0.5$. That is,

$$\log\left(\frac{1/2}{1-1/2}\right) = 0 = \alpha + \beta \log(\text{LD50})$$

hence,

$$\text{LD50} = \exp(-\alpha/\beta).$$

(b) What is the MLE, $\widehat{\text{LD50}}$, of LD50?

Solution. In general, the MLE of $\widehat{g(\boldsymbol{\theta})}$, some function of unknow parameters, is just $g(\hat{\boldsymbol{\theta}})$. We have therefore $\widehat{\text{LL50}} = \exp(-\hat{\alpha}/\hat{\beta})$.

Problem 3.

(a) How would you define the risk of a woman giving birth to a child with a birth defect, after exposure to contaminated drinking water during pregnancy?

Solution. Let

$$\begin{aligned} A &= \{\text{woman giving birth to a child with a birth defec}\}, \\ B &= \{\text{exposure to contaminated drinking water during pregnancy}\} \end{aligned}$$

be the two events in question. Then,

$$\text{risk} = \Pr(A \mid B).$$

(b) How would you define the relative risk of a birth defect, where "relative" here refers to exposure versus no exposure?

Solution. Denote by A^c and B^c the *complements* of event A and B, respectively. Then

$$\text{relative risk} = \frac{\Pr(A \mid B)}{\Pr(A \mid B^c)}.$$

(c) Same question as in b), except replace "relative risk" with "odds ratio".

Solution.

$$\text{odds ratio} = \frac{\Pr(A \mid B)/\Pr(A^c \mid B)}{Pr(A \mid B^c)/\Pr(A^c \mid B^c)}.$$

Problem 4. Suppose a measurement X is made of a true concentration C. The "Method LOD" is defined as,

$$\text{LOD} = k \cdot \text{var}(X \mid C = 0),$$

Recall that the distribution of $(X \mid C = 0)$ is lognormal (μ, σ^2) if and only if

$$(Y \mid C = 0) \sim \text{normal}(\mu, \sigma^2),$$

where $Y = \log X$. Suppose we want to be sure that $\Pr(X \geq \text{LOD} \mid C = 0) = 2.5\%$. Assuming that X is lognormally distributed, find a formula for k in terms of μ and σ, where you will need to know that the upper 2.5% point for a standard normal distribution is 1.96 and,

$$\text{var}(X \mid C = 0) = \exp(2\mu + \sigma^2)(\exp(\sigma^2) - 1).$$

Solution. First, using $\Pr(X \geq \text{LOD} \mid C = 0) = 2.5\%$, solve for LOD in terms of μ and σ. Now,

$$\Pr(X \geq \text{LOD} \mid C = 0) = 2.5\%$$

is equivalent to

$$\Pr\left(\frac{\log(X) - \mu}{\sigma} \geq \frac{\log(\text{LOD}) - \mu}{\sigma} \,\middle|\, C = 0\right) = 2.5\%, \tag{1}$$

since the logarithmic function is monotonically increasing. But, $(\log(X) - \mu)/\sigma \sim \text{normal}(0, 1)$ and the 97.5% (= 100% − 2.5%) quantile of the standard normal distribution is 1.96. That is, (1) holds if

$$\frac{\log(\text{LOD}) - \mu}{\sigma} = 1.96,$$

hence,

$$\text{LOD} = \exp(\mu + 1.96\sigma).$$

Finally, from $\text{LOD} = k \text{ var}(X \mid C = 0)$, we get

$$k = \frac{\text{LOD}}{\text{var}(X \mid C = 0)} = \frac{\exp(\mu + 1.96\sigma)}{\exp(2\mu + \sigma^2)(\exp(\sigma^2) - 1)} = \frac{\exp(-\mu + 1.96\sigma - \sigma^2)}{\exp(\sigma^2) - 1}.$$