STA 6351, Report.1.2

Carson Slater *Baylor University*

1.2

This hypothetical model is motivated by settings where a continuous outcome is observed only if a preliminary count variable meets a threshold.

- For example, in a clinical trial, let Y be a Poisson-distributed immune cell count, such as for neutrophils, which target bacteria and fungi. Suppose that in this trial, such counts are available for all subjects, but an expensive downstream assay (yielding a continuous response X, such as cytokine concentration) is performed only if the immune response is low ($Y \leq c$) for some positive integer c.
- Similarly, in an industrial reliability study, let Y denote the number of defects in a manufactured item. If the defect count exceeds a screening threshold, the item is scrapped and no further testing occurs; otherwise ($Y \leq c$) the item undergoes a lifetime or strength test recorded as X.
- In both examples, the observed data consist of *all* Poisson counts, together with normal outcomes on a selective subsample. This requires constructing the likelihood by combining contributions from both sources of information.

Here is the model setup. Let $c \in \mathbb{Z}_{\geq 0}$ be fixed. For each unit $i = 1, \dots, n$, we observe a Poisson count

$$Y_i \sim \text{Pois}(\mu),$$

and we observe a continuous outcome

$$X_i \sim \text{N}(\mu, \sigma^2)$$

only if $Y_i \leq c$. Define the indicator $\Delta_i \equiv \mathbb{I}\{Y_i \leq c\}$. Thus the recorded data are

$$\mathcal{D} = \{(Y_i, \Delta_i, X_i^{\text{obs}}) : i = 1, \dots, n\}, \quad X_i^{\text{obs}} = \begin{cases} X_i, & \Delta_i = 1, \\ \text{missing}, & \Delta_i = 0. \end{cases}$$

Assume that $Y_i \perp X_i \mid \mu, \sigma^2$ and that subjects are IID given (μ, σ^2). Unless otherwise stated, treat $\sigma^2 > 0$ and c as known.

Carry out the following:

(a)

For a *single* unit, write the joint model for $(Y, \Delta, X^{\text{obs}})$. Show that, for any realized y and $\delta = \mathbb{I}\{y \leq c\}$,

$$P_{Y,\Delta,X^{\text{obs}}}(y, \delta, x) = \Pr(Y = y) \times \underbrace{\left[\frac{1}{\sigma}\varphi\left(\frac{x-\mu}{\sigma}\right)\right]^{\delta}}_{\text{Normal if } \delta=1},$$

where φ is the standard normal PDF, and the bracketed factor is interpreted as 1 when $\delta = 0$ (no x is observed). Briefly justify why it is valid to omit any additional factor for the "missing" X when $Y > c$.

For a single unit, we need to write the joint model for $(Y, \Delta, X^{\text{obs}})$ where:

- $Y \sim \text{Pois}(\mu)$ (always observed)
- $\Delta = \mathbb{I}\{Y \leq c\}$ (deterministic given Y)
- $X^{\text{obs}} = X$ if $\Delta = 1$, missing if $\Delta = 0$
- $X \sim \text{N}(\mu, \sigma^2)$ when observed

Since Δ is a deterministic function of Y, we have $\Delta = \delta$ if and only if the constraint $\{Y \leq c\} = \{\delta = 1\}$ or $\{Y > c\} = \{\delta = 0\}$ is satisfied.

Case 1: $\delta = 1$ **(i.e.,** $y \leq c$**)**

In this case, we observe both $Y = y$ and $X = x$. Since $Y \perp X \mid \mu, \sigma^2$, the joint density is:

$$P_{Y,\Delta,X^{\text{obs}}}(y, 1, x) = p_Y(y) \cdot p_X(x) = \frac{e^{-\mu}\mu^y}{y!} \cdot \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

This can be written as:

$$P_{Y,\Delta,X^{\text{obs}}}(y, 1, x) = \Pr(Y = y) \times \frac{1}{\sigma}\varphi\left(\frac{x-\mu}{\sigma}\right)$$

Case 2: $\delta = 0$ **(i.e.,** $y > c$**)**

In this case, we observe only $Y = y$ and X is missing. The joint density is simply the marginal density of Y:

$$P_{Y,\Delta,X^{\text{obs}}}(y, 0, \cdot) = P_Y(y) = \Pr(Y = y)$$

Combining both cases, we can write:

$$P_{Y,\Delta,X^{\text{obs}}}(y, \delta, x) = \Pr(Y = y) \times \left[\frac{1}{\sigma}\varphi\left(\frac{x-\mu}{\sigma}\right)\right]^{\delta}$$

where the bracketed factor equals 1 when $\delta = 0$ (since any number raised to the power 0 equals 1).

When $Y > c$ (i.e., $\delta = 0$), the value of X is not observed due to the study design, not due to any random mechanism. This is a form of *missing completely at random* (MCAR) conditional on Y. Since the missingness is deterministic given the observed Y, we don't need to model the probability of missingness separately - it's already captured by $P(Y = y)$ for $y > c$. The likelihood contribution from unit i when $\delta_i = 0$ is simply the probability of observing $Y_i = y_i$, with no additional terms needed for the unobserved X_i.

(b)

Using independence across i, derive the full-sample likelihood $L(\mu|\mathcal{D})$ (with σ^2 treated as known). Express your result as a product of a Poisson part and a normal part:

$$L(\mu|\mathcal{D}) = \prod_{i=1}^{n} \Pr(Y_i = y_i) \times \prod_{i:\Delta_i=1} \frac{1}{\sigma}\varphi\left(\frac{x_i - \mu}{\sigma}\right).$$

(Hint: The observation rule is fully determined by Y_i, which is observed; the resulting missingness is therefore *ignorable* for likelihood construction.) Identify the contributions coming from the $\{Y_i\}$ and from the observed $\{X_i : \Delta_i = 1\}$.

Given the independence assumption, the likelihood is the product of individual unit contributions:

$$L(\mu|\mathcal{D}) = \prod_{i=1}^{n} P_{Y_i,\Delta_i,X_i^{\text{obs}}}(y_i, \delta_i, x_i)$$

From part (a), we know that for each unit i:

$$P_{Y_i,\Delta_i,X_i^{\text{obs}}}(y_i, \delta_i, x_i) = P(Y_i = y_i) \times \left[\frac{1}{\sigma}\varphi\left(\frac{x_i - \mu}{\sigma}\right)\right]^{\delta_i}$$

Substituting into the likelihood:

$$L(\mu|\mathcal{D}) = \prod_{i=1}^{n} P(Y_i = y_i) \times \prod_{i=1}^{n} \left[\frac{1}{\sigma}\varphi\left(\frac{x_i - \mu}{\sigma}\right)\right]^{\delta_i}$$

Since $Y_i \sim \text{Pois}(\mu)$, we have $P(Y_i = y_i) = \frac{e^{-\mu}\mu^{y_i}}{y_i!}$.

For the second product, note that the factor $\left[\frac{1}{\sigma}\varphi\left(\frac{x_i-\mu}{\sigma}\right)\right]^{\delta_i}$ equals:

- $\frac{1}{\sigma}\varphi\left(\frac{x_i-\mu}{\sigma}\right)$ when $\delta_i = 1$ (i.e., X_i is observed)
- 1 when $\delta_i = 0$ (i.e., X_i is missing)

Therefore, the second product simplifies to:

$$\prod_{i=1}^{n} \left[\frac{1}{\sigma}\varphi\left(\frac{x_i - \mu}{\sigma}\right)\right]^{\delta_i} = \prod_{i:\delta_i=1} \frac{1}{\sigma}\varphi\left(\frac{x_i - \mu}{\sigma}\right)$$

So the final likelihood expression is:

$$\begin{aligned} L(\mu|\mathcal{D}) &= \prod_{i=1}^{n} \frac{e^{-\mu}\mu^{y_i}}{y_i!} \times \prod_{i:\delta_i=1} \frac{1}{\sigma}\varphi\left(\frac{x_i-\mu}{\sigma}\right) \\ &= \frac{e^{-n\mu}\mu^{\sum_{i=1}^{n} y_i}}{\prod_{i=1}^{n} y_i!} \times \prod_{i:\delta_i=1} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x_i-\mu)^2}{2\sigma^2}\right) \end{aligned}$$

We identify the contributions to the likelihood as follows:

Poisson part (from the $\{Y_i\}$):

$$L_{\text{Poisson}}(\mu) = \prod_{i=1}^{n} P(Y_i = y_i) = \frac{e^{-n\mu}\mu^{\sum_{i=1}^{n} y_i}}{\prod_{i=1}^{n} y_i!}$$

This depends on μ through the sufficient statistics n (total number of observations) and $\sum_{i=1}^{n} y_i$ (total count across all units).

Normal part (from the observed $\{X_i : \delta_i = 1\}$):

$$L_{\text{Normal}}(\mu) = \prod_{i:\delta_i=1} \frac{1}{\sigma}\varphi\left(\frac{x_i-\mu}{\sigma}\right)$$

This depends on μ through the observed X_i values for units where $\delta_i = 1$ (i.e., where $Y_i \leq c$).

The key insight from the hint is that since the observation rule is fully determined by the observed Y_i values, the missingness pattern is *ignorable* for likelihood construction - we don't need additional terms to account for the missing data mechanism.