

Report on Application of Logistic functions to Bio-Assay by Joseph Berkson

Parth Agrawal

October 2025

1 Introduction

In this paper, Joseph Berkson talks about the Bio-Assay which is nothing more than determining the potency of a substance like a drug or a poison. We measure all-or-nothing biological effects, such as mortality rates, as a function of dose.

1.1 The Dose-Response Curve and Logarithmic Scale

When we plot the effect (q) vs. the log of dose (x), we obtain an S-shaped curve or a **sigmoidal curve**.

Normally, we try to explain it as an **integral of the normal distribution**, which stems from the fact that susceptibility to the drug is normally distributed.

However, Berkson proposes that we should use the Logistic Function instead due to its already wide applications in physicochemical phenomena.

2 The Logistic Function and the Logit Transformation

2.1 Mathematical Definitions

The logistic function for the estimated mortality rate (\hat{Q}) is defined by the non-linear equation:

$$\hat{Q} = \frac{1}{1 + e^{-(\alpha + \beta x)}} \quad \text{or} \quad \hat{P} = \frac{e^{-(\alpha + \beta x)}}{1 + e^{-(\alpha + \beta x)}} \quad (\text{where } \hat{P} = 1 - \hat{Q} \text{ is the survival rate})$$

Here, α and β are the parameters (constants) that define the S-curve. $\hat{Q} = 0.5$ (or LD_{50}) occurs when $\alpha + \beta x = 0$, giving the relationship $\text{Log LD}_{50} = -\alpha/\beta$.

2.2 The Logit Transformation

The **Logit** of the observed mortality rate (q) is defined as the natural logarithm of the ratio of the observed mortality rate to the observed survival rate ($p = 1 - q$):

$$\text{Logit}(q) \equiv y = \ln\left(\frac{q}{p}\right) = \ln\left(\frac{q}{1-q}\right)$$

Applying this transformation to the logistic function yields a **linear relationship** (a powerful simplification familiar from linearizing non-linear functions like $\ln(y) = mx + c$):

$$y = \alpha + \beta x$$

This linear form is the foundation of the simplified statistical method similar to applying a logitc regression in classification problems.

3 Statistical Methods: Least Squares vs. Maximum Likelihood

Berkson wants to use the logistic function because it allows him to implement the **Method of Least Squares**.

3.1 Least Squares on the Linearized Logit

The goal of the LS method is to find parameters ($\hat{\alpha}, \hat{\beta}$) that minimize the sum of weighted squared errors. For the linear logit model ($y = \alpha + \beta x$), the standard approach is complicated because the variance of the logit ($\text{Var}(y)$) depends on the unknown, fitted parameters.

However, Berkson shows that by a suitable approximation the weights (ω_i) required for the Weighted Least Squares (WLS) can be formulated **entirely in terms of the observed quantities**:

$$\text{Weight: } \omega_i = n_i \cdot q_i \cdot p_i$$

where n_i is the number of individuals exposed at dose x_i , q_i is the observed mortality rate, and p_i is the observed survival rate.

Since weights can be obtained independently of fitted parameters $\hat{\alpha}$ and $\hat{\beta}$, we can obtain a solution directly in one step without using linear regression formulas.

3.2 Comparison with Maximum Likelihood

The competing method, **Maximum Likelihood** (used with the Integrated Normal Curve), aims to choose parameters that maximize the probability of observing the given data. This typically involves using calculus to find the stationary points of the log-likelihood function (\mathcal{L}):

$$\frac{\partial \mathcal{L}}{\partial \alpha} = 0 \quad \text{and} \quad \frac{\partial \mathcal{L}}{\partial \beta} = 0$$

Berkson explicitly critiques the ML method as being complex and potentially yielding **biased estimates** with **larger standard errors** compared to his proposed WLS solution.

4 Conclusion and Comparative Findings

Berkson’s empirical comparison (Table II) using several real-world bio-assay datasets validates his proposal.

- **Goodness-of-Fit (χ^2):** The comparison of the sum of weighted squared deviations (χ^2) showed that the fit provided by the Logistic Function was **practically the same as or better than** the Integrated Normal Curve, with no instances where the normal curve proved superior.
- **Final Thesis:** Berkson successfully demonstrates that the Logistic Function, when coupled with the linearizing Logit Transformation, provides a **statistically simpler** and equally effective method for bio-assay analysis, making the Weighted Least Squares approach highly preferable to the iterative Maximum Likelihood method for this specific class of problems.