

Pseudo-Poisson Distributions with Nonlinear Conditional Rates

Jared N. Lakhani

Department of Statistical Sciences, University of Cape Town

lkhjar001@myuct.ac.za

Abstract

Arnold and Manjunath [1] claim that the bivariate pseudo-Poisson distribution is well suited to bivariate count data with one equidispersed and one overdispersed marginal, owing to its parsimonious structure and straightforward parameter estimation. In the formulation of Leiter and Hamdan [2], the conditional mean of X_2 was specified as a function of X_1 ; Arnold and Manjunath [1] subsequently augmented this specification by adding an intercept, yielding a linear conditional rate. A direct implication of this construction is that the bivariate pseudo-Poisson distribution can represent only positive correlation between the two variables. This study generalizes the conditional rate to accommodate negatively correlated datasets by introducing curvature. This augmentation provides the additional benefit of allowing the model to behave approximately linear when appropriate, while adequately handling the boundary case $(x_1, x_2) = (0, 0)$. According to the Akaike Information Criterion (AIC), the models proposed in this study outperform Arnold and Manjunath [1]'s linear models.

1 Introduction

We review the concept of the bivariate pseudo-Poisson distribution as discussed in Arnold and Manjunath [1].

Definition 1. *A pair of variables (X_1, X_2) is said to have a bivariate pseudo-Poisson distribution if there exists a positive constant λ_1 such that:*

$$X_1 \sim \text{Poisson}(\lambda_1),$$

and a function $\lambda_2 : \{0, 1, 2, \dots\} \rightarrow (0, \infty)$ such that, for every non-negative integer x_1 :

$$X_2 \mid X_1 \sim \text{Poisson}(\lambda_2(x_1)).$$

Motivated by Example 1 in Arnold and Manjunath [3], we specify the conditional rate as $\lambda_2(x_1) = \delta + \beta F(x_1; \boldsymbol{\theta})$, where $\delta \geq 0$ is an intercept, $\beta \in \mathbb{R}$ is an amplitude parameter, and $F(x_1; \boldsymbol{\theta})$ is a parametric distribution function on $[0, \infty)$ (with $\lambda_2(x_1) \geq 0$ for all x_1). This parameterization permits negative correlation between X_1 and X_2 when $\beta < 0$ and $\delta \geq \max\{-\beta, 0\}$ to ensure $\lambda_2(x_1) \geq 0$. By Theorem 1, the sign of $\text{Corr}(X_1, X_2)$ is determined by the monotonicity of $\lambda_2(x_1)$. Since $F(x_1; \boldsymbol{\theta})$ is always non-decreasing, the monotonicity of $\lambda_2(x_1)$ is dependent on the sign of β . Hence, if λ_2 is strictly increasing in x_1 ($\beta > 0$), the correlation is positive; if λ_2 is strictly decreasing ($\beta < 0$), the correlation is negative.

Theorem 1. *Given a pseudo-Poisson form $X_1 \sim \text{Poisson}(\lambda_1)$, $X_2 \mid X_1 \sim \text{Poisson}(\lambda_2(x_1))$ with $\lambda_1, \lambda_2(x_1) > 0$ for $x_1, x_2 \in \{0, 1, 2, \dots\}$: if $\lambda_2(x_1)$ is strictly increasing with respect to x_1 , then $\rho(X_1, X_2) > 0$.*

Proof. We note $E(X_1) = \alpha$, $E(X_2) = E_{X_1}\{E_{X_2|X_1}(X_2 | X_1 = x_1)\} = E_{X_1}\{\lambda_2(X_1)\}$. Furthermore, $E(X_1 X_2) = E_{X_1}\{X_1 E_{X_2|X_1}(X_2 | X_1 = x_1)\} = E_{X_1}\{X_1 \lambda_2(X_1)\}$. Now we have:

$$\begin{aligned}
Cov(X_1, X_2) &= E(X_1 X_2) - E(X_1)E(X_2) \\
&= E_{X_1}\{X_1 \lambda_2(X_1)\} - \lambda_1 E_{X_1}\{\lambda_2(X_1)\} \\
&= \sum_{i=0}^{\infty} i \lambda_2(i) \frac{e^{-\lambda_1} \lambda_1^i}{i!} - \sum_{i=0}^{\infty} \lambda_1 \lambda_2(i) \frac{e^{-\lambda_1} \lambda_1^i}{i!} \\
&= \sum_{i=1}^{\infty} i \lambda_2(i) \frac{e^{-\lambda_1} \lambda_1^i}{i!} - \sum_{i=0}^{\infty} \lambda_1 \lambda_2(i) \frac{e^{-\lambda_1} \lambda_1^i}{i!} \\
&= \sum_{i=0}^{\infty} (i+1) \lambda_2(i+1) \frac{e^{-\lambda_1} \lambda_1^{i+1}}{(i+1)!} - \sum_{i=0}^{\infty} \lambda_1 \lambda_2(i) \frac{e^{-\lambda_1} \lambda_1^i}{i!} \\
&= \sum_{i=0}^{\infty} \lambda_1 \lambda_2(i+1) \frac{e^{-\lambda_1} \lambda_1^i}{i!} - \sum_{i=0}^{\infty} \lambda_1 \lambda_2(i) \frac{e^{-\lambda_1} \lambda_1^i}{i!} \\
&= \sum_{i=0}^{\infty} (i+1) \frac{e^{-\lambda_1} \lambda_1^i}{(i+1)!} (\lambda_2(i+1) - \lambda_2(i))
\end{aligned}$$

Since $\lambda_2(i+1) > \lambda_2(i)$, it follows that $Cov(X_1, X_2) > 0$, and consequently $\rho(X_1, X_2) > 0$. Conversely, if $\lambda_2(x_1)$ is strictly decreasing such that $\lambda_2(i+1) < \lambda_2(i)$, then $Cov(X_1, X_2) < 0$, implying $\rho(X_1, X_2) < 0$. When $\lambda_2(x_1)$ is constant, $\rho(X_1, X_2) = 0$. \square

Additionally, we offer Theorem 2, merely to show an alternative way of proving Theorem 6.3 in Arnold and Manjunath [1], claiming the bivariate pseudo-Poisson distribution is always over-dispersed.

Theorem 2. Given a pseudo-Poisson form $X_1 \sim Poisson(\lambda_1)$, $X_2 | X_1 \sim Poisson(\lambda_2(x_1))$, $Var(X_1) = E(X_1)$ and $Var(X_2) \geq E(X_2)$ given $\lambda_1, \lambda_2(x_1) > 0$ for $x_1, x_2 \in \{0, 1, 2, \dots\}$.

Proof. The first identity follows directly from the properties of the Poisson distribution. Since $X_1 \sim Poisson(\lambda_1)$, we have $Var(X_1) = E(X_1) = \lambda_1$. Using $Var_{X_2|X_1}(X_2|X_1) = E_{X_2|X_1}(X_2|X_1) - E(X_2|X_1)^2$ from the properties of a Poisson distribution since $X_2 | X_1$ is Poisson distributed, we have:

$$\begin{aligned}
Var(X_2) &= E_{X_1}\{Var_{X_2|X_1}(X_2|X_1)\} + Var_{X_1}\{E_{X_2|X_1}(X_2|X_1)\} \\
&= E_{X_1}\{E_{X_2|X_1}(X_2|X_1)\} + Var_{X_1}\{E_{X_2|X_1}(X_2|X_1)\} \\
&= E(X_2) + Var_{X_1}\{\lambda_2(X_1)\} \\
&\geq E(X_2)
\end{aligned}$$

where the inequality follows since variance is always non-negative. \square

In this study, the distribution function $F(x_1; \theta)$ is specified to belong to either the exponential or Lomax family. Importantly, these two were chosen because closed-form expressions for their moments can be obtained, with infinite sums simplified either via the exponential Maclaurin series or by using hypergeometric functions. The study further segregates these full models - models with full parameterization - into sub-models in which certain parameters are nullified (by fixing them at 0 or 1), inspired by Arnold and Manjunath [1]. The study specifies method-of-moments estimators (m.m.e.'s) either through implicit relations or explicitly stated equations, examines the correlations each model can accommodate, and includes likelihood-ratio tests to compare the sub-models with their respective full models. Additionally, simulated data are provided for enrichment, after which two datasets (also used

in Arnold and Manjunath [1]) are analysed as applications of the models. Furthermore, in terms of model fit according to the Akaike Information Criterion (AIC), this study's models outperform Arnold and Manjunath [1]'s linear models. We attribute this to the curvature of the conditional rate $\lambda_2(x_1)$, where this extends to their added ability to model $(x_1, x_2) = (0, 0)$ observations adequately. This study's models also allow for negatively correlated bivariate datasets - although no such data is analysed here.

2 Exponential

We restrict $F(x_1; \theta)$, with $\theta = \gamma$, to be the exponential distribution function. Thus, we have:

$$X_1 \sim \text{Poisson}(\alpha) \quad (1)$$

$$X_2 | X_1 \sim \text{Poisson} (\delta + \beta (1 - e^{-\gamma x_1})) , \quad (2)$$

where the rate parameter $\gamma > 0$ governs the rate of change (gradient) of the regression function. Later, we re-parameterize via $e^{-\gamma} = \nu$. The parameters satisfy $\alpha, \gamma > 0$. To ensure that $\lambda_2(x_1) > 0$ for all $x_1 > 0$, it is necessary that $\beta \in \mathbb{R} \setminus \{0\}$ with $\delta \geq \max\{-\beta, 0\}$ - which corresponds to $\delta \geq -\beta$ in the case of negative correlation ($\beta < 0$) and $\delta \geq 0$ when the correlation is positive ($\beta > 0$). We ensure $\beta \neq 0$ here so to ensure an independence model is not created - that is, where X_2 is no longer conditioned on X_1 .

We also examine various sub-models that arise from the exponential model by constraining specific parameters. These include removing the amplitude control - thereby fixing positive and negative correlation by setting $\beta = 1$ and $\beta = -1$ respectively, fixing the rate parameter to $\gamma = 1$, or jointly constraining both. In addition, we consider a no-intercept model by setting $\delta = 0$, which by construction precludes the possibility of negative correlation. We further note, that for the no-intercept model, Case III: $\delta = 0$, when $x_1 = 0$, the conditional distribution of x_2 degenerates at zero, implying that $x_2 = 0$ with probability one. Consequently, any empirical observation with $x_1 = 0$ and $x_2 > 0$ is assigned zero probability under the model, which yields a likelihood of zero and indicates model misspecification for such observations. Figure 4 provides a schematic illustration of the resulting sub-models.

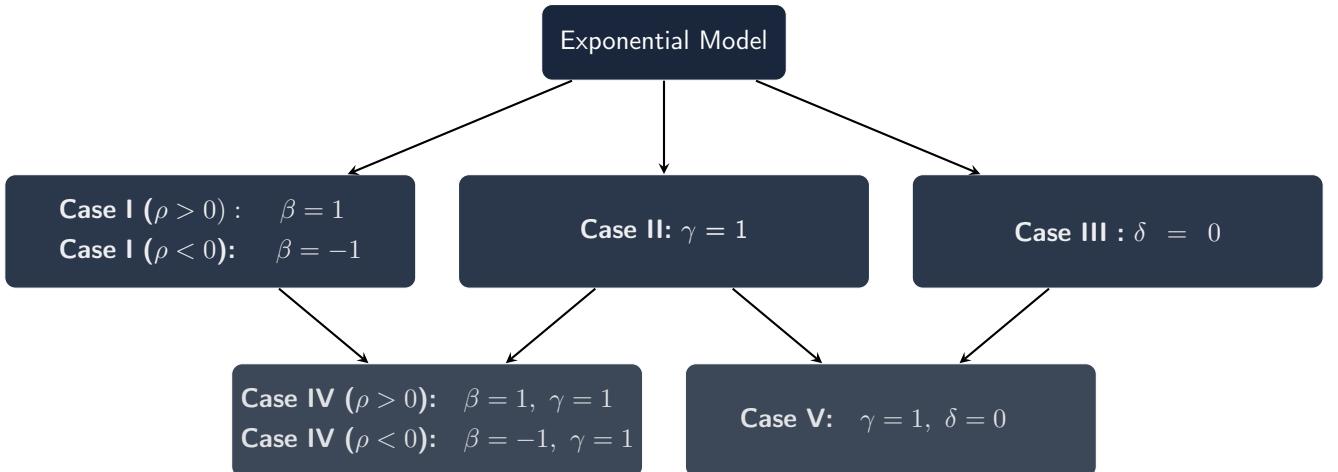


Figure 1: Flow diagram of exponential sub-models.

2.1 Probability mass functions

Following from the definition of a Poisson distribution:

$$\begin{aligned}
P(X_1) &= \frac{e^{-\alpha} \alpha^{x_1}}{x_1!}, \\
P(X_2|X_1) &= \frac{e^{-\delta-\beta(1-e^{-\gamma x_1})} \{\delta + \beta(1 - e^{-\gamma x_1})\}^{x_2}}{x_2!}, \\
P(X_1, X_2) &= P(X_1)P(X_2|X_1) \\
&= \frac{e^{-\alpha} \alpha^{x_1}}{x_1!} \frac{e^{-\delta-\beta(1-e^{-\gamma x_1})} \{\delta + \beta(1 - e^{-\gamma x_1})\}^{x_2}}{x_2!},
\end{aligned} \tag{3}$$

where $x_1, x_2 \in \{0, 1, 2, \dots\}$, $\alpha, \gamma > 0$ and $\beta \in \mathbb{R} \setminus \{0\}$ with $\delta \geq \max(-\beta, 0)$.

2.2 Moments

After substituting $e^\alpha = \mu$ and $e^{-\gamma} = \nu$, we note:

$$E(X_1) = \alpha, \tag{4}$$

$$\begin{aligned}
E(X_2) = E_{X_1}\{E(X_2|X_1 = x_1)\} &= \delta + \beta - \beta E_{X_1}(\nu^{X_1}) \\
&= \delta + \beta(1 - \mu^{\nu-1}),
\end{aligned} \tag{5}$$

Since

$$\begin{aligned}
E_{X_1}(\nu^{X_1}) &= \sum_{i=0}^{\infty} \nu^i \frac{\mu^{-1} \alpha^i}{i!} \\
&= \mu^{-1} \sum_{i=0}^{\infty} \frac{(\alpha \nu)^i}{i!} \\
&= \mu^{-1} e^{\alpha \nu} \\
&= \mu^{\nu-1}
\end{aligned} \tag{6}$$

We note, since $0 < \nu < 1 \implies \frac{1}{\mu} < \mu^{\nu-1} < 1$ confirming $E(X_2) > 0$. Now

$$Var(X_1) = \alpha, \tag{7}$$

$$\begin{aligned}
Var_{X_2|X_1}(X_2|X_1) &= \delta + \beta(1 - \nu^{X_1}) \\
&= E_{X_2|X_1}(X_2|X_1),
\end{aligned} \tag{8}$$

Hence, from Equations 5, 6, 8:

$$\begin{aligned}
Var(X_2) &= E_{X_1}\{Var_{X_2|X_1}(X_2|X_1)\} + Var_{X_1}\{E_{X_2|X_1}(X_2|X_1)\} \\
&= E(X_2) + \beta^2 Var_{X_1}(\nu^{X_1}) \\
&= E(X_2) + \beta^2 \{E_{X_1}(\nu^{2X_1}) - \{E_{X_1}(\nu^{X_1})\}^2\} \\
&= \beta(1 - \mu^{\nu-1}) + \delta + \beta^2(\mu^{\nu^2-1} - \mu^{2\nu-2}).
\end{aligned} \tag{9}$$

Note $0 < \nu < 1 \implies \mu^{\nu^2-1} > \mu^{2\nu-2} \implies \text{Var}(X_2) > E(X_2)$, which is a stricter version of Theorem 2. Furthermore from Equation 4:

$$\begin{aligned}
E(X_1 X_2) = E_{X_1}\{E(X_1 X_2 | X_1)\} &= E_{X_1}\{X_1 E(X_2 | X_1)\} \\
&= E_{X_1}\{X_1 (\delta + \beta(1 - e^{-\gamma X_1}))\} \\
&= \alpha(\delta + \beta) - \beta E_{X_1}(X_1 \nu^{X_1}) \\
&= \alpha(\delta + \beta(1 - \nu \mu^{\nu-1})),
\end{aligned} \tag{10}$$

since

$$\begin{aligned}
E_{X_1}(X_1 \nu^{X_1}) &= \sum_{i=0}^{\infty} i \nu^i \frac{\mu^{-1} \alpha^i}{i!} \\
&= \mu^{-1} \sum_{i=1}^{\infty} \alpha \nu \frac{(\alpha \nu)^{(i-1)}}{(i-1)!} \\
&= \mu^{-1} \alpha \nu \sum_{n=0}^{\infty} \frac{(\alpha \nu)^n}{n!} \\
&= \mu^{-1} \alpha \nu e^{\alpha \nu} \\
&= \alpha \nu \mu^{\nu-1}.
\end{aligned}$$

Now from Equations 4, 5, 10, the covariance between X_1 and X_2 is:

$$\begin{aligned}
\text{Cov}(X_1, X_2) &= E(X_1 X_2) - E(X_1)E(X_2) \\
&= \alpha \beta (1 - \nu) \mu^{\nu-1}.
\end{aligned} \tag{11}$$

We note that since $0 < \nu < 1 \implies \alpha(1 - \nu) \mu^{\nu-1} > 0$. Consequently, the sign of the covariance - and thus the direction of the correlation between X_1 and X_2 - is determined solely by the sign of β , as previously noted. Now the associated correlation using Equations 7, 9 and 11:

$$\begin{aligned}
\rho(\alpha, \beta, \gamma, \delta) &= \frac{\text{Cov}(X_1 X_2)}{\sqrt{\text{Var}(X_1) \text{Var}(X_2)}} \\
&= \frac{\sqrt{\alpha} \beta (1 - \nu) \mu^{\nu-1}}{\sqrt{\beta^2 (\mu^{\nu^2-1} - \mu^{2\nu-2}) + \beta(1 - \mu^{\nu-1}) + \delta}}.
\end{aligned} \tag{12}$$

2.2.1 Method of moments estimation

Suppose data in the form of $\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \dots, \mathbf{X}^{(n)}$ (where for each $i = 1, \dots, n$, $\mathbf{X}^{(i)} = (X_{1i}, X_{2i})^T$) are obtained, which are independent and identically distributed according to Equations 1 - 2. If $M_1, M_2 > 0$, consistent asymptotically normal method of moment estimates (m.m.e's) are acquired.

Define:

$$\begin{aligned} M_1 &= \frac{1}{n} \sum_{i=1}^n x_{1i}, \\ M_2 &= \frac{1}{n} \sum_{i=1}^n x_{2i}, \\ S_{12} &= \frac{1}{n} \sum_{i=1}^n (x_{1i} - M_1)(x_{2i} - M_2), \\ S_{22} &= \frac{1}{n} \sum_{i=1}^n (x_{2i} - M_2)^2. \end{aligned}$$

Now equating Equation 4 to its sample mean:

$$\begin{aligned} \tilde{\alpha} &= M_1, \\ e^{\tilde{\alpha}} &= e^{M_1} = \tilde{\mu}, \end{aligned} \tag{13}$$

where Equation 13 is the applicable m.m.e for all models. Consider the remaining moments, $\tilde{\beta}$, $\tilde{\gamma}$ and $\tilde{\delta}$ for each model separately:

2.2.1.1 Full exponential model:

Taking the quotient of Equation 11 squared and the difference between Equation 9 and Equation 5, and equating the result to the solution obtained from identical operations on the associated sample moments, an implicit equation w.r.t $\tilde{\nu}$ (hence $\tilde{\gamma}$) is obtained, to be solved numerically:

$$\frac{\tilde{\alpha}^2(\tilde{\nu}-1)^2}{\tilde{\mu}^{(\tilde{\nu}-1)^2}-1} = \frac{S_{12}^2}{S_2 - M_2}, \tag{14}$$

and to ensure the expression remains real-valued we obtain the condition:

$$S_2 \neq M_2,$$

although we already elucidated that $Var(X_2) > E(X_2)$ for $\beta \neq 0$. Since $\forall x_1, x_2 \in \{0, 1, 2, \dots\}$, $\tilde{\gamma} > 0 \implies 0 < \tilde{\nu} < 1$. Hence, since Equation 14 is strictly increasing for $0 < \tilde{\nu} < 1$ from Theorem 3 (see Appendix 7 for proof), we may use $\lim_{\tilde{\nu} \rightarrow 0^+}$ and $\lim_{\tilde{\nu} \rightarrow 1^-}$ to obtain the lower and upper bounds, respectively, of the left side of Equation 14, hence $\frac{S_{12}^2}{S_2 - M_2}$. Therefore:

$$\frac{M_1^2}{e^{M_1} - 1} < \frac{S_{12}^2}{S_2 - M_2} < M_1. \tag{15}$$

Figure 2 illustrates $\tilde{\nu}$ as a function of $\tilde{\alpha} = M_1$ and $\frac{S_{12}^2}{S_2 - M_2}$. Importantly, due to the bounds imposed by M_1 on $\frac{S_{12}^2}{S_2 - M_2}$ in Equation 15, certain combinations of M_1 , S_2 , M_2 and S_{12} values may lead to the non-existence of $\tilde{\gamma}$ (hence $\tilde{\nu}$). This behavior is also evident from the patterns observed in Figure 2.

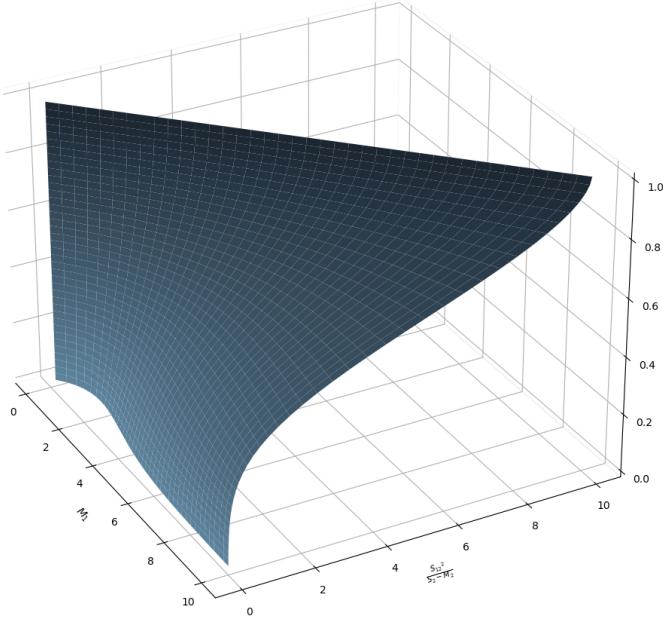


Figure 2: $\tilde{\nu}$ vs M_1 and $\frac{S_{12}^2}{S_2 - M_2}$ as per Equation 14.

We then solve for $\tilde{\beta}$ by equating Equation 11 to the sample covariance to obtain:

$$\tilde{\beta} = \frac{S_{12}}{\tilde{\alpha}(1 - \tilde{\nu})\tilde{\mu}^{\tilde{\nu}-1}}. \quad (16)$$

Finally, solve for $\tilde{\delta}$ by merely equating Equation 5 to the sample mean:

$$\tilde{\delta} = M_2 - \tilde{\beta}(1 - \tilde{\mu}^{\tilde{\nu}-1}). \quad (17)$$

We note, that the m.m.e's of Case 3: $\delta = 0$ may be derived analogously to the approach previously outlined, with the exception of excluding the estimator for δ . Additionally, we observe that Equation 14 may be rewritten as $\frac{\zeta}{e^\zeta - 1} = \frac{S_{12}^2}{M_1(S_2 - M_2)}$, where $\zeta = M_1(\tilde{\nu} - 1)^2$. The function $\frac{\zeta}{e^\zeta - 1}$ is well-known in mathematical analysis and appears in the generating function for the Bernoulli numbers B_n , satisfying the expansion $\frac{\zeta}{e^\zeta - 1} = \sum_{n=0}^{\infty} \frac{B_n \zeta^n}{n!} = 1 - \frac{\zeta}{2} + \frac{\zeta^2}{12} - \frac{\zeta^4}{720} + \dots$. This series representation may offer a more computationally efficient approach for approximating $\tilde{\nu}$ in the regime of small $\zeta = M_1(\tilde{\nu} - 1)^2$.

2.2.1.2 Case I ($\beta = 1$ or $\beta = -1$):

Solving for $\tilde{\gamma}$ (and thus $\tilde{\nu}$), one may either apply Equation 14, or alternatively, equate the model-based expression for $Cov(X_1, X_2)$ (Equation 11) with the sample covariance. Using the latter, this yields:

$$\tilde{\alpha}(1 - \tilde{\nu})\tilde{\mu}^{\tilde{\nu}-1}\text{sign}(\beta) = S_{12}.$$

Applying the Lambert W function allows us to express $\tilde{\nu}$ explicitly as:

$$\tilde{\nu} = W\left(-\frac{S_{12}}{\tilde{\alpha}}\text{sign}(\beta)\right) + 1.$$

Naturally, since the Lambert W function is only defined for arguments satisfying $-\frac{S_{12}}{\tilde{\alpha}}\text{sign}(\beta) = -\frac{S_{12}}{M_1}\text{sign}(\beta) \geq -\frac{1}{e}$, a necessary condition for the existence of a real-valued solution is $\frac{M_1}{S_{12}}\text{sign}(\beta) \geq e$. Furthermore, since $0 < \tilde{\nu} < 1$, we restrict attention to the

principal branch W_0 of the Lambert W function to ensure that the solution remains within the admissible parameter space, with the additional requirement that $-\frac{S_{12}}{\tilde{\alpha}} \text{sign}(\beta) < 0 \implies \text{sign}(\beta)S_{12} > 0$.

Furthermore, we solve for $\tilde{\delta}$ just as before, by merely equating Equation 5 to the sample mean:

$$\tilde{\delta} = M_2 - \text{sign}(\beta)(1 - \tilde{\mu}^{\tilde{\nu}-1}).$$

2.2.1.3 Case II ($\gamma = 1$):

Since $\gamma = 1$, we have $\nu = \frac{1}{e}$, hence we equate the model-based expression for $Cov(X_1, X_2)$ (Equation 11) with the sample covariance to obtain:

$$\tilde{\beta} = \frac{S_{12}}{\tilde{\alpha}(1 - \frac{1}{e})\tilde{\mu}^{(\frac{1}{e}-1)}}.$$

Afterwhich, we solve for $\tilde{\delta}$ equating Equation 5 to the sample mean:

$$\tilde{\delta} = M_2 - \tilde{\beta}(1 - \tilde{\mu}^{\frac{1}{e}-1}).$$

2.2.1.4 Case IV ($\gamma = 1$ and $\beta = 1$ or $\gamma = 1$ and $\beta = -1$):

We solve for $\tilde{\delta}$ by merely equating Equation 5 to the sample mean:

$$\tilde{\delta} = M_2 - \text{sign}(\beta)(1 - \tilde{\mu}^{\frac{1}{e}-1}).$$

2.2.1.5 Case V ($\delta = 0$ and $\gamma = 1$):

We solve for $\tilde{\beta}$ by merely equating Equation 5 to the sample mean:

$$\tilde{\beta} = \frac{M_2}{1 - \tilde{\mu}^{\frac{1}{e}-1}}.$$

2.3 Correlations

We analyse the correlation between X_1 and X_2 . Because each model imposes bounds on the attainable correlation ρ , we report these feasible ranges explicitly; otherwise users might apply a model to data with ρ outside its admissible set, leading to misspecification or nonexistence of estimators. We utilise the multivariate ρ expression in Equation 12, that is $\rho(\alpha, \beta, \gamma, \delta)$ noting again $\mu = e^\alpha$ and $\nu = e^{-\nu}$. Now one may simply ascertain ρ bounds through numerical optimisation (and we utilise this as a safety check as well) to find said bounds, yet we use the approach of studying asymptotic regimes - where these limits reveal which directions in the parameter space push ρ toward its attainable extrema. Furthermore, we note that $\lim_{\alpha \rightarrow 0^+} \rho(\alpha, \beta, \gamma, \delta) = 0$ (since X_1 becomes nearly degenerate and the induced dependence vanishes), hence we do not take this limit first in the subsequent sections as this would not offer us greater insight.

2.3.1 Full exponential model

With regards to the full exponential model, we note that by taking the limits of Equation 12(limits are commutable):

$$\lim_{\gamma \rightarrow \infty} \lim_{\beta \rightarrow \infty} \rho(\alpha, \beta, \gamma, \delta) = \sqrt{\frac{\alpha}{\mu - 1}}, \quad (18)$$

and

$$\lim_{\gamma \rightarrow \infty} \lim_{\beta \rightarrow -\infty} \rho(\alpha, \beta, \gamma, \delta) = -\sqrt{\frac{\alpha}{\mu - 1}}. \quad (19)$$

Seeing as these results are not functions of δ , we confirm that the bounds of the correlation function of the full exponential model, $\rho(\alpha, \beta, \gamma, \delta)$, and the bounds of the correlation function of the Case III sub-model, $\rho(\alpha, \beta, \gamma)$, are identical. These bounds being $-1 < \rho < 1$ by taking $\lim_{\alpha \rightarrow 0^+}$ of Equation 18 and 19 noting said equations are monotonic with respect to α .

2.3.2 Case I

2.3.2.1 $\beta = 1$

Similarly, we take the limits of Equation 12 after substituting in $\beta = 1$, and noting for positive correlation (that is, $\beta > 0$) we must have $\delta \geq 0$, we obtain (limits are commutable):

$$\lim_{\gamma \rightarrow \infty} \lim_{\delta \rightarrow 0^+} \rho(\alpha, \beta = 1, \gamma, \delta) = \sqrt{\frac{\alpha}{\mu^2 - 1}}, \quad (20)$$

which elucidates that the bounds for ρ are $0 < \rho < \frac{\sqrt{2}}{2}$ after taking $\lim_{\alpha \rightarrow 0}$ to Equation 20, noting said equations monotonicity with respect to α .

2.3.2.2 $\beta = -1$

Likewise, we take the limits of Equation 12 after substituting in $\beta = -1$, and noting for negative correlation we must have $\delta \geq -\beta$, we obtain (limits are commutable):

$$\lim_{\gamma \rightarrow \infty} \lim_{\delta \rightarrow 1^+} \rho(\alpha, \beta = -1, \gamma, \delta) = -\sqrt{\frac{\alpha}{2\mu - 1}}, \quad (21)$$

Since Equation 21 is not monotonic with respect to α , we rather computationally minimise: $\alpha^* = \operatorname{argmin}_\alpha -\sqrt{\frac{\alpha}{2\mu - 1}} = 0.76804$, such that $-\sqrt{\frac{\alpha^*}{2e^{\alpha^*} - 1}} = -0.48162$. Hence, we obtain the bounds $-0.48162 < \rho < 0$.

2.4 Case II

Similarly, taking the limits of Equation 12:

$$\lim_{\beta \rightarrow \infty} \rho(\alpha, \beta, \gamma = 1, \delta) = -\frac{\sqrt{\alpha} \frac{1-e}{e}}{\sqrt{e^{\frac{\alpha(1-e)^2}{e^2}} - 1}}, \quad (22)$$

and

$$\lim_{\beta \rightarrow -\infty} \rho(\alpha, \beta, \gamma = 1, \delta) = \frac{\sqrt{\alpha} \frac{1-e}{e}}{\sqrt{e^{\frac{\alpha(1-e)^2}{e^2}} - 1}}. \quad (23)$$

After which, taking $\lim_{\alpha \rightarrow 0^+}$ on Equation 22 and 23, we obtain $-1 < \rho < 1$.

2.5 Case IV

2.5.0.1 $\beta = 1$

Again, taking the limits of Equation 12, and noting for positive correlation (that is, $\beta > 0$) we must have $\delta \geq 0$:

$$\lim_{\delta \rightarrow 0^+} \rho(\alpha, \beta = 1, \gamma = 1, \delta) = - \frac{\sqrt{\alpha} \frac{1-e}{e} e^{\frac{\alpha}{e}}}{\sqrt{e^{2\alpha} \left(1 - e^{2\alpha \frac{1-e}{e}}\right) - e^{\frac{\alpha}{e} + \alpha} \left(e^{\alpha \frac{1-e}{e^2}} - 1\right)}} \quad (24)$$

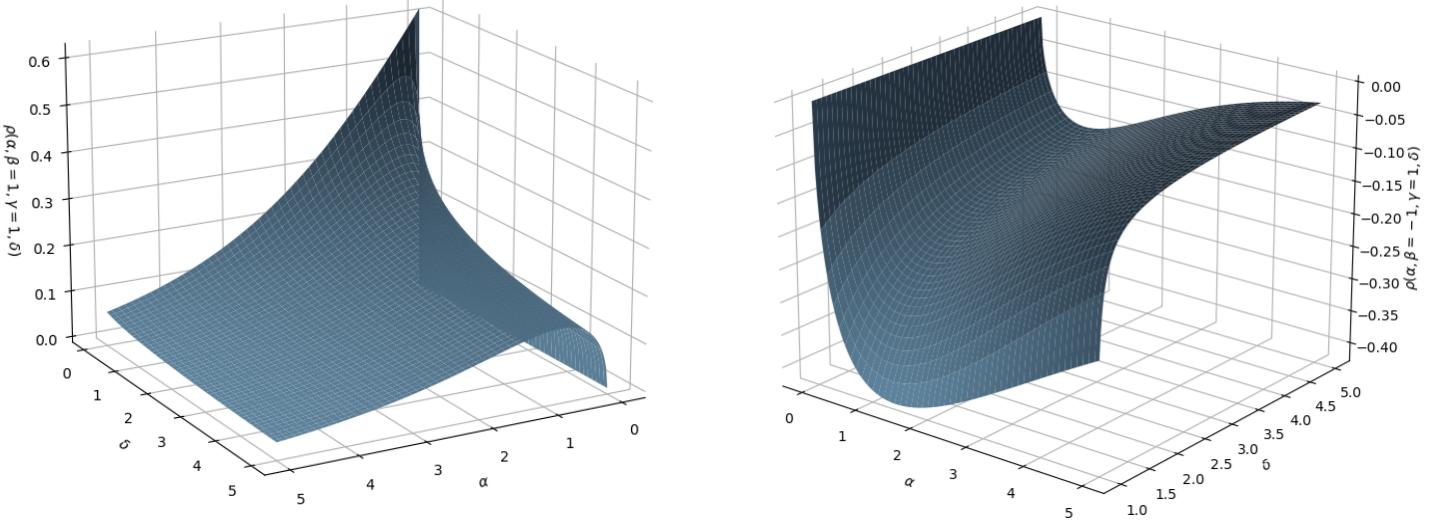
Noting the monotonicity of Equation 24, we take $\lim_{\alpha \rightarrow 0^+}$ to obtain $\sqrt{\frac{e-1}{2e-1}} \approx 0.62233$. Hence, our bounds are $0 < \rho < \sqrt{\frac{e-1}{2e-1}} \approx 0.62233$ also illustrated in Figure 3a displaying $\rho(\alpha, \beta = 1, \gamma = 1, \delta)$.

2.5.0.2 $\beta = -1$

Again, taking the limits of Equation 12, and noting for negative correlation (that is, $\beta < 0$) we must have $\delta \geq -\beta$:

$$\lim_{\delta \rightarrow 1^+} \rho(\alpha, \beta = 1, \gamma = 1, \delta) = \frac{\sqrt{\alpha} \frac{1-e}{e} e^{\frac{\alpha}{e}}}{\sqrt{e^{\frac{\alpha}{e} + \alpha} \left(e^{\alpha \frac{1-e}{e^2}} - e^{\alpha \frac{1-e}{e}} + 1\right)}} \quad (25)$$

Now Equation 25 is not monotonic with respect to α , so instead of taking limits, we rather minimise computationally: $\alpha^* = \operatorname{argmin}_\alpha \{\lim_{\delta \rightarrow 1^+} \rho(\alpha, \beta = 1, \gamma = 1, \delta)\} = 1.427885$, such that $\{\lim_{\delta \rightarrow 1^+} \rho(\alpha, \beta = 1, \gamma = 1, \delta)\}_{|\alpha=\alpha^*} \approx -0.419946$. Hence, our bounds are $\approx -0.419946 < \rho < 0$ also illustrated in Figure 3b displaying $\rho(\alpha, \beta = -1, \gamma = 1, \delta)$.



(a) $\rho(\alpha, \beta = 1, \gamma = 1, \delta)$ against $\alpha > 0$ and $\delta \geq 0$ noting a maximum achieved at ≈ 0.622 .

(b) $\rho(\alpha, \beta = -1, \gamma = 1, \delta)$ against $\alpha > 0$, $\delta \geq 1$ noting a minimum achieved at ≈ -0.42 .

Figure 3: $\rho(\alpha, \beta = 1, \gamma = 1, \delta)$ and $\rho(\alpha, \beta = -1, \gamma = 1, \delta)$.

2.6 Case V

Similarly, taking the limits of Equation 12 (limits are commutable):

$$\lim_{\alpha \rightarrow 0^+} \lim_{\beta \rightarrow \infty} \rho(\alpha, \beta, \gamma = 1, \delta = 0) = 1,$$

and

$$\lim_{\alpha \rightarrow 0^+} \lim_{\beta \rightarrow -\infty} \rho(\alpha, \beta, \gamma = 1, \delta = 0) = -1.$$

Hence, the bounds are $-1 < \rho < 1$.

We summarise the correlation $\rho(\alpha, \beta, \gamma, \delta)$ bounds in Table 1, confirmed with numerical optimisation.

Model	ρ	Lower Bound	Upper Bound
Full Exponential (including Case III)		-1	1
Case I ($\beta = 1$)		0	$\frac{\sqrt{2}}{2} \approx 0.707$
Case I ($\beta = -1$)		≈ -0.482	0
Case II ($\gamma = 1$)		-1	1
Case IV ($\beta = 1, \gamma = 1$)		0	$\sqrt{\frac{e-1}{2e-1}} \approx 0.622$
Case IV ($\beta = -1, \gamma = 1$)		≈ -0.420	0
Case V ($\delta = 0, \gamma = 1$)		-1	1

Table 1: ρ lower and upper bounds for the full exponential model and sub-models.

2.7 Maximum likelihood estimation

Consider again, the data in the form of $\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \dots, \mathbf{X}^{(n)}$ (where for each $i = 1, \dots, n$, $\mathbf{X}^{(i)} = (X_{1i}, X_{2i})^T$) are obtained, which are independent and identically distributed according to Equations 1 - 2, the the likelihood function is as follows from Equation 3, utilising $\nu = e^{-\gamma}$:

$$L(\alpha, \beta, \gamma, \delta) = \prod_{i=1}^n \frac{e^{-\alpha} \alpha^{x_{1i}} e^{-(\delta + \beta(1 - e^{-\gamma x_{1i}}))} (\delta + \beta(1 - e^{-\gamma x_{1i}}))^{x_{2i}}}{x_{1i}! x_{2i}!} \\ = \frac{e^{-n(\alpha+\beta+\delta)} \alpha^{\sum_{i=1}^n x_{1i}} e^{\beta \sum_{i=1}^n \nu^{x_{1i}}} \prod_{i=1}^n (\delta + \beta(1 - \nu^{x_{1i}}))^{x_{2i}}}{\prod_{i=1}^n (x_{1i}! x_{2i}!)}$$

Now the log-likelihood is:

$$l = \log(L(\alpha, \beta, \gamma, \delta)) = -n(\alpha + \beta + \delta) + \log(\alpha) \sum_{i=1}^n x_{1i} + \beta \sum_{i=1}^n \nu^{x_{1i}} + \sum_{i=1}^n x_{2i} \log(\delta + \beta(1 - \nu^{x_{1i}})) \\ + h(\mathbf{x}_1, \mathbf{x}_2), \quad (26)$$

where $h(\mathbf{x}_1, \mathbf{x}_2) = -\log(\prod_{i=1}^n (x_{1i}! x_{2i}!))$. Now since $\frac{\partial l}{\partial \alpha} = 0 \implies \hat{\alpha} = \frac{\sum_{i=1}^n x_{1i}}{n} = \tilde{\alpha}$, that is, the m.l.e of α is equivalent to the m.m.e of α . The m.l.e's of β, γ, δ are to be solved numerically using Equation 26

2.8 Likelihood ratio tests

The general form of the likelihood ratio test is:

$$\Lambda = \frac{\sup_{\boldsymbol{\theta} \in \Theta_0} L(\boldsymbol{\theta})}{\sup_{\boldsymbol{\theta} \in \Theta} L(\boldsymbol{\theta})}. \quad (27)$$

Under regularity conditions, if H_0 holds, then $-2 \log(\Lambda) \xrightarrow{d} \chi_{\dim(\Theta) - \dim(\Theta_0)}^2$. We reject H_0 when $-2 \log(\Lambda) > \chi_{\dim(\Theta) - \dim(\Theta_0), 1-\alpha}^2$ where $\chi_{\dim(\Theta) - \dim(\Theta_0), 1-\alpha}^2$ denotes the $(1 - \alpha)$ critical value of the chi-square distribution with $\dim(\Theta) - \dim(\Theta_0)$ degrees of freedom. We note the natural parameter space of the full exponential model is $\Theta = \{(\alpha, \beta, \gamma, \delta)^T : \alpha > 0, \gamma > 0, \beta \in \mathbb{R} \setminus \{0\}, \delta \geq \max(-\beta, 0)\}$. Let $\hat{\alpha}, \hat{\beta}, \hat{\gamma}, \hat{\delta}$ be the m.l.e's of the full exponential model. Furthermore, we denote $\hat{\alpha}^*, \hat{\beta}^*, \hat{\gamma}^*, \hat{\delta}^*$ as any m.l.e under H_0 in this section.

2.8.1 Testing $\beta = \pm 1$

Under H_0 , the natural parameter space is $\Theta_0 = \{(\alpha, \gamma, \delta)^T : \alpha > 0, \gamma > 0, \delta \geq \mathbb{I}(\beta < 0)\}$. The generalized likelihood ratio test statistic in Equation 27 is:

$$\Lambda = \frac{\frac{e^{-n(\hat{\alpha}^* + \text{sign}(\beta) + \hat{\delta}^*)} \hat{\alpha}^{*\sum_{i=1}^n x_{1i}} e^{\text{sign}(\beta) \sum_{i=1}^n (\hat{\nu}^*)^{x_{1i}}} \prod_{i=1}^n (\hat{\delta}^* + \text{sign}(\beta)(1 - (\hat{\nu}^*)^{x_{1i}}))^{x_{2i}}}{\prod_{i=1}^n (x_{1i}! x_{2i}!)}}{\frac{e^{-n(\hat{\alpha} + \hat{\beta} + \hat{\delta})} \hat{\alpha}^{\sum_{i=1}^n x_{1i}} e^{\hat{\beta} \sum_{i=1}^n \hat{\nu}^{x_{1i}}} \prod_{i=1}^n (\hat{\delta} + \hat{\beta}(1 - \hat{\nu}^{x_{1i}}))^{x_{2i}}}{\prod_{i=1}^n (x_{1i}! x_{2i}!)}}.$$

Since $\hat{\alpha}^* = \hat{\alpha}$, we have:

$$\Lambda = e^{-n(\text{sign}(\beta) + \hat{\delta}^* - \hat{\beta} - \hat{\delta})} e^{(\text{sign}(\beta) \sum_{i=1}^n (\hat{\nu}^*)^{x_{1i}} - \hat{\beta} \sum_{i=1}^n \hat{\nu}^{x_{1i}})} \prod_{i=1}^n \left[\frac{\hat{\delta}^* + \text{sign}(\beta)(1 - (\hat{\nu}^*)^{x_{1i}})}{\hat{\delta} + \hat{\beta}(1 - \hat{\nu}^{x_{1i}})} \right]^{x_{2i}}.$$

And by taking the logarithm, we obtain:

$$\begin{aligned} \log \Lambda = & -n \left(\text{sign}(\beta) + \hat{\delta}^* - \hat{\beta} - \hat{\delta} \right) + \text{sign}(\beta) \sum_{i=1}^n (\hat{\nu}^*)^{x_{1i}} - \hat{\beta} \sum_{i=1}^n \hat{\nu}^{x_{1i}} \\ & + \sum_{i=1}^n x_{2i} \log \left[\frac{\hat{\delta}^* + \text{sign}(\beta)(1 - (\hat{\nu}^*)^{x_{1i}})}{\hat{\delta} + \hat{\beta}(1 - \hat{\nu}^{x_{1i}})} \right]. \end{aligned}$$

2.8.2 Testing $\gamma = 1$

Under H_0 , the natural parameter space is $\Theta_0 = \{(\alpha, \beta, \delta)^T : \alpha > 0, \beta \in \mathbb{R} \setminus \{0\}, \delta \geq \max(-\beta, 0)\}$. The logarithm of the generalized likelihood ratio test statistic in Equation 27 is:

$$\log \Lambda = -n \left(\hat{\beta}^* + \hat{\delta}^* - \hat{\beta} - \hat{\delta} \right) + \hat{\beta}^* \sum_{i=1}^n \left(\frac{1}{e} \right)^{x_{1i}} - \hat{\beta} \sum_{i=1}^n \hat{\nu}^{x_{1i}} + \sum_{i=1}^n x_{2i} \log \left[\frac{\hat{\delta}^* + \hat{\beta}^*(1 - (\frac{1}{e})^{x_{1i}})}{\hat{\delta} + \hat{\beta}(1 - \hat{\nu}^{x_{1i}})} \right].$$

2.8.3 Testing $\delta = 0$

Under H_0 , the natural parameter space is $\Theta_0 = \{(\alpha, \beta, \gamma)^T : \alpha > 0, \beta \in \mathbb{R} \setminus \{0\}, \gamma > 0\}$. The logarithm of the generalized likelihood ratio test statistic in Equation 27 is:

$$\log \Lambda = -n \left(\hat{\beta}^* - \hat{\beta} - \hat{\delta} \right) + \hat{\beta}^* \sum_{i=1}^n (\hat{\nu}^*)^{x_{1i}} - \hat{\beta} \sum_{i=1}^n \hat{\nu}^{x_{1i}} + \sum_{i=1}^n x_{2i} \log \left[\frac{\hat{\beta}^*(1 - (\hat{\nu}^*)^{x_{1i}})}{\hat{\delta} + \hat{\beta}(1 - \hat{\nu}^{x_{1i}})} \right].$$

2.8.4 Testing $\beta = \pm 1$ and $\gamma = 1$

Under H_0 , the natural parameter space is $\Theta_0 = \{(\alpha, \delta)^T : \alpha > 0, \delta \geq \mathbb{I}(\beta < 0)\}$. The logarithm of the generalized likelihood ratio test statistic in Equation 27 is:

$$\log \Lambda = -n \left(\text{sign}(\beta) + \hat{\delta}^* - \hat{\beta} - \hat{\delta} \right) + \text{sign}(\beta) \sum_{i=1}^n \left(\frac{1}{e} \right)^{x_{1i}} - \hat{\beta} \sum_{i=1}^n \hat{\nu}^{x_{1i}} + \sum_{i=1}^n x_{2i} \log \left[\frac{\hat{\delta}^* + \text{sign}(\beta)(1 - (\frac{1}{e})^{x_{1i}})}{\hat{\delta} + \hat{\beta}(1 - \hat{\nu}^{x_{1i}})} \right].$$

2.8.5 Testing $\delta = 0$ and $\gamma = 1$

Under H_0 , the natural parameter space is $\Theta_0 = \{(\alpha, \beta)^T : \alpha > 0, \beta \in \mathbb{R} \setminus \{0\}\}$. The logarithm of the generalized likelihood ratio test statistic in Equation 27 is:

$$\log \Lambda = -n \left(\hat{\beta}^* - \hat{\beta} - \hat{\delta} \right) + \hat{\beta}^* \sum_{i=1}^n \left(\frac{1}{e} \right)^{x_{1i}} - \hat{\beta} \sum_{i=1}^n \hat{\nu}^{x_{1i}} + \sum_{i=1}^n x_{2i} \log \left[\frac{\hat{\beta}^* \left(1 - \left(\frac{1}{e} \right)^{x_{1i}} \right)}{\hat{\delta} + \hat{\beta} \left(1 - \hat{\nu}^{x_{1i}} \right)} \right].$$

3 Lomax

We restrict $F(x_1; \theta)$, with $\theta = [\gamma', \eta']^T$, to be the Lomax distribution function. Thus, we have:

$$X_1 \sim \text{Poisson}(\alpha') \quad (28)$$

$$X_2 | X_1 \sim \text{Poisson} \left(\delta' + \beta' \left(1 - \left(\frac{\gamma'}{x_1 + \gamma'} \right)^{\eta'} \right) \right), \quad (29)$$

where the scale parameter $\gamma' > 0$ and shape parameter $\eta' > 0$ govern the rate of change (gradient) of the regression function. The parameters satisfy $\alpha', \gamma', \eta' > 0$. To ensure that $\lambda_2(x_1) > 0$ for all $x_1 > 0$, it is necessary that $\beta' \in \mathbb{R} \setminus \{0\}$ with $\delta' \geq \max(-\beta', 0)$ - which corresponds to $\delta' \geq -\beta'$ in the case of negative correlation ($\beta' < 0$) and $\delta' \geq 0$ when the correlation is positive ($\beta' > 0$).

We also examine various sub-models that arise from the $\text{Lomax}(\eta' = 1, \gamma')$ model by constraining specific parameters. These include removing the amplitude control - thereby fixing positive and negative correlation by setting $\beta' = 1$ and $\beta' = -1$ respectively, fixing the rate parameter as $\gamma' = 1$, or jointly constraining both. In addition, we consider a no-intercept model by setting $\delta' = 0$, which by construction precludes the possibility of negative correlation. We further note (as before with the full exponential model), that for the no-intercept model, Case III: $\delta' = 0$, when $x_1 = 0$, the conditional distribution of x_2 degenerates at zero, implying that $x_2 = 0$ with probability one. Consequently, any empirical observation with $x_1 = 0$ and $x_2 > 0$ is assigned zero probability under the model, which yields a likelihood of zero and indicates model misspecification for such observations. Figure 4 provides a schematic illustration of the resulting sub-models.

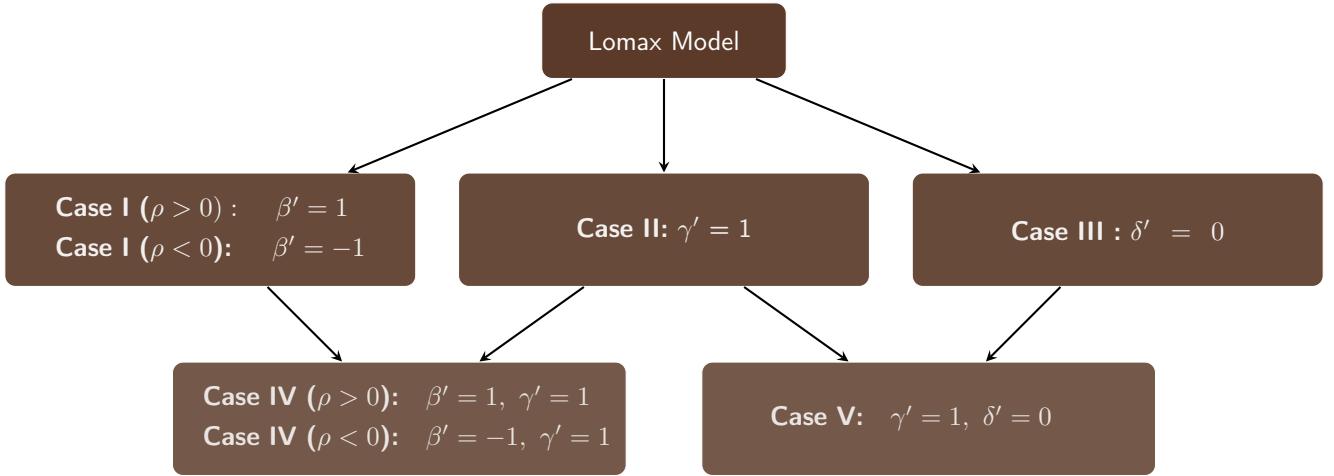


Figure 4: Flow diagram of Lomax sub-models.

Before proceeding, we note that, for the exponential model, taking the limit $\lim_{\gamma \rightarrow \infty}$ of the regression form in Equation 2

yields the same result as taking $\lim_{\gamma' \rightarrow 0^+}$ of the regression form in Equation 29. Conversely, taking the limit $\lim_{\gamma \rightarrow 0^+}$ in Equation 2 corresponds to $\lim_{\gamma' \rightarrow \infty}$ in Equation 29. We summarize these equivalences below:

$$\begin{aligned}\lim_{\gamma \rightarrow \infty} \{\delta + \beta (1 - e^{-\gamma x_1})\} &= \delta + \beta, \\ \lim_{\gamma' \rightarrow 0^+} \left\{ \delta' + \beta' \left(1 - \left(\frac{\gamma'}{x_1 + \gamma'} \right)^{\eta'} \right) \right\} &= \delta' + \beta',\end{aligned}$$

and

$$\begin{aligned}\lim_{\gamma \rightarrow 0} \{\delta + \beta (1 - e^{-\gamma x_1})\} &= \delta, \\ \lim_{\gamma' \rightarrow \infty} \left\{ \delta' + \beta' \left(1 - \left(\frac{\gamma'}{x_1 + \gamma'} \right)^{\eta'} \right) \right\} &= \delta'.\end{aligned}$$

This implies that under these limits, the exponential and Lomax models become equivalent, with $\alpha = \alpha'$, $\beta = \beta'$ and $\delta = \delta'$ (where η' becomes inconsequential under these limits). Specifically, we have:

$$X_1 \sim \text{Poisson}(\alpha), \quad X_2 | X_1 \sim \text{Poisson}(\delta + \beta) \quad \text{as } \lim_{\gamma \rightarrow \infty} \text{ or } \lim_{\gamma' \rightarrow 0^+},$$

and

$$X_1 \sim \text{Poisson}(\alpha), \quad X_2 | X_1 \sim \text{Poisson}(\delta) \quad \text{as } \lim_{\gamma \rightarrow 0} \text{ or } \lim_{\gamma' \rightarrow \infty}.$$

That is to say, both exponential and Lomax models become independence models - where X_2 is no longer conditioned on X_1 .

3.1 Probability mass functions

Following from the definition of a Poisson distribution:

$$\begin{aligned}P(X_1) &= \frac{e^{-\alpha'} (\alpha')^{x_1}}{x_1!}, \\ P(X_2|X_1) &= \frac{e^{-\delta' - \beta' \left(1 - \left(\frac{\gamma'}{x_1 + \gamma'} \right)^{\eta'} \right)} \left\{ \delta' + \beta' \left(1 - \left(\frac{\gamma'}{x_1 + \gamma'} \right)^{\eta'} \right) \right\}^{x_2}}{x_2!}, \\ P(X_1, X_2) &= P(X_1)P(X_2|X_1) \\ &= \frac{e^{-\alpha'} (\alpha')^{x_1}}{x_1!} \frac{e^{-\delta' - \beta' \left(1 - \left(\frac{\gamma'}{x_1 + \gamma'} \right)^{\eta'} \right)} \left\{ \delta' + \beta' \left(1 - \left(\frac{\gamma'}{x_1 + \gamma'} \right)^{\eta'} \right) \right\}^{x_2}}{x_2!},\end{aligned} \tag{30}$$

where $x_1, x_2 \in \{0, 1, 2, \dots\}$, $\alpha', \gamma', \eta' > 0$ and $\beta' \in \mathbb{R} \setminus \{0\}$ with $\delta' \geq \max(-\beta', 0)$.

3.2 Moments

We again denote $e^{\alpha'} = \mu'$, and for ease of simplification, we express the summations in terms of hypergeometric functions, and accordingly restrict η' to take discrete values; that is, we assume $\eta' \in \mathbb{N}$. Now:

$$E(X_1) = \alpha', \quad (31)$$

$$\begin{aligned} E(X_2) &= E_{X_1}\{E(X_2|X_1=x_1)\} = \delta' + \beta' - \beta'E_{X_1}\left\{\left(\frac{\gamma'}{x_1+\gamma'}\right)^{\eta'}\right\} \\ &= \delta' + \beta'\left(1 - \frac{1}{\mu'} {}_{\eta'}F_{\eta'}(\gamma', \dots, \gamma', (\gamma'+1), \dots, (\gamma'+1); \alpha')\right), \end{aligned} \quad (32)$$

since:

$$\begin{aligned} E_{X_1}\left\{\left(\frac{\gamma'}{x_1+\gamma'}\right)^{\eta'}\right\} &= \sum_{i=0}^{\infty} \left(\frac{\gamma'}{\gamma'+i}\right)^{\eta'} \frac{e^{-\alpha'}(\alpha')^i}{i!} \\ &= \sum_{i=0}^{\infty} \left(\frac{(\gamma')_i}{(\gamma'+1)_i}\right)^{\eta'} \frac{e^{-\alpha'}(\alpha')^i}{i!} \\ &= \frac{1}{e^{\alpha'} \eta'} {}_{\eta'}F_{\eta'}(\gamma', \dots, \gamma', (\gamma'+1), \dots, (\gamma'+1); \alpha'), \end{aligned} \quad (33)$$

and using the Pochhammer symbol $(q)_n = \begin{cases} 1 & \text{if } n = 0 \\ q(q+1)\cdots(q+n-1) & \text{if } n > 0, \end{cases}$

$$\begin{aligned} \frac{\gamma'}{\gamma'+i} &= \frac{\gamma'(\gamma'+1)(\gamma'+2)\cdots(\gamma'+i-2)(\gamma'+i-1)}{(\gamma'+1)(\gamma'+2)(\gamma'+3)\cdots(\gamma'+i-1)(\gamma'+i)} \\ &= \frac{(\gamma')_i}{(\gamma'+1)_i}. \end{aligned}$$

We denote ${}_{\eta'}F_{\eta'}(\gamma', \dots, \gamma', (\gamma'+1), \dots, (\gamma'+1); \alpha')$ as ${}_{\eta'}F_{\eta'}(\gamma')$ going forward. Now,

$$Var(X_1) = \alpha', \quad (34)$$

$$\begin{aligned} Var_{X_2|X_1}(X_2 | X_1) &= \delta' + \beta'\left(1 - \left(\frac{\gamma'}{x_1+\gamma'}\right)^{\eta'}\right) \\ &= E_{X_2|X_1}(X_2 | X_1). \end{aligned} \quad (35)$$

Using Equations 32, 33 and 35, we have:

$$\begin{aligned} Var(X_2) &= E_{X_1}\{Var_{X_2|X_1}(X_2|X_1)\} + Var_{X_1}\{E_{X_2|X_1}(X_2|X_1)\} \\ &= E(X_2) + (\beta')^2 Var_{X_1}\left\{\left(\frac{\gamma'}{x_1+\gamma'}\right)^{\eta'}\right\} \\ &= E(X_2) + (\beta')^2 \left(E_{X_1}\left\{\left(\frac{\gamma'}{x_1+\gamma'}\right)^{2\eta'}\right\} - \left(E_{X_1}\left\{\left(\frac{\gamma'}{x_1+\gamma'}\right)^{\eta'}\right\}\right)^2\right) \\ &= \delta' + \beta'\left(1 - \frac{1}{\mu'} {}_{\eta'}F_{\eta'}(\gamma')\right) + (\beta')^2 \left(\frac{1}{\mu'} {}_{2\eta'}F_{2\eta'}(\gamma') - \frac{1}{(\mu')^2} ({}_{\eta'}F_{\eta'}(\gamma'))^2\right). \end{aligned} \quad (36)$$

Furthermore from Equation 31:

$$\begin{aligned}
E(X_1 X_2) &= E_{X_1}\{E(X_1 X_2 | X_1)\} = E_{X_1}\{X_1 E(X_2 | X_1)\} \\
&= E_{X_1}\left\{X_1 \left(\delta' + \beta' \left(1 - \left(\frac{\gamma'}{x_1 + \gamma'}\right)^{\eta'}\right)\right)\right\} \\
&= \alpha'(\delta' + \beta') - \beta' E_{X_1}\left\{X_1 \left(\frac{\gamma'}{x_1 + \gamma'}\right)^{\eta'}\right\} \\
&= \alpha' \left(\delta' + \beta' \left(1 - \frac{1}{\mu'} \left(\frac{\gamma'}{\gamma' + 1}\right)^{\eta'} {}_{\eta'}F_{\eta'}(\gamma' + 1)\right)\right). \tag{37}
\end{aligned}$$

since:

$$\begin{aligned}
E_{X_1}\left\{X_1 \left(\frac{\gamma'}{x_1 + \gamma'}\right)^{\eta'}\right\} &= \sum_{i=0}^{\infty} i \left(\frac{\gamma'}{\gamma' + i}\right)^{\eta'} \frac{e^{-\alpha'} (\alpha')^i}{i!} \\
&= \sum_{i=1}^{\infty} \left(\frac{\gamma'}{\gamma' + (i-1) + 1}\right)^{\eta'} \frac{e^{-\alpha'} (\alpha')^{i-1} \alpha'}{(i-1)!} \\
&= \alpha' \frac{1}{e^{\alpha'}} \left(\frac{\gamma'}{\gamma' + 1}\right)^{\eta'} \sum_{j=0}^{\infty} \left(\frac{\gamma' + 1}{\gamma' + j + 1}\right)^{\eta'} \frac{(\alpha')^j}{j!} \\
&= \alpha' \frac{1}{e^{\alpha'}} \left(\frac{\gamma'}{\gamma' + 1}\right)^{\eta'} \sum_{j=0}^{\infty} \left(\frac{(\gamma' + 1)_j}{(\gamma' + 2)_j}\right)^{\eta'} \frac{(\alpha')^j}{j!} \\
&= \alpha' \frac{1}{e^{\alpha'}} \left(\frac{\gamma'}{\gamma' + 1}\right)^{\eta'} {}_{\eta'}F_{\eta'}((\gamma' + 1), \dots, (\gamma' + 1), (\gamma' + 2), \dots, (\gamma' + 2); \alpha').
\end{aligned}$$

and using:

$$\begin{aligned}
\frac{\gamma' + 1}{\gamma' + j + 1} &= \frac{(\gamma' + 1)(\gamma' + 2)(\gamma' + 3) \cdots (\gamma' + 1 + j - 2)(\gamma' + 1 + j - 1)}{(\gamma' + 2)(\gamma' + 3)(\gamma' + 4) \cdots (\gamma' + 2 + j - 2)(\gamma' + 2 + j - 1)} \\
&= \frac{(\gamma' + 1)_j}{(\gamma' + 2)_j}.
\end{aligned}$$

Now from Equations 31, 32 and 37, the covariance between X_1 and X_2 is:

$$\begin{aligned}
Cov(X_1, X_2) &= E(X_1 X_2) - E(X_1)E(X_2) \\
&= \frac{\alpha' \beta'}{\mu'} \left({}_{\eta'}F_{\eta'}(\gamma') - \left(\frac{\gamma'}{\gamma' + 1}\right)^{\eta'} {}_{\eta'}F_{\eta'}(\gamma' + 1)\right). \tag{38}
\end{aligned}$$

We note that since ${}_{\eta'}F_{\eta'}(\gamma') = \sum_{i=1}^{\infty} \left(\frac{\gamma'}{\gamma' + i}\right)^{\eta'} \frac{\alpha'}{i!}$ and $\left(\frac{\gamma'}{\gamma' + 1}\right)^{\eta'} {}_{\eta'}F_{\eta'}(\gamma' + 1) = \sum_{i=1}^{\infty} \left(\frac{\gamma'}{\gamma' + 1 + i}\right)^{\eta'} \frac{\alpha'}{i!}$, we confirm that ${}_{\eta'}F_{\eta'}(\gamma') > \left(\frac{\gamma'}{\gamma' + 1}\right)^{\eta'} {}_{\eta'}F_{\eta'}(\gamma' + 1)$ implying that the sign of $Cov(X_1, X_2)$ is solely dictated by the sign of β' as postulated before. Now the associated correlation between X_1 and X_2 , using Equations 34, 36 and 38, is:

$$\begin{aligned}
\rho(\alpha', \beta', \gamma', \delta', \eta') &= \frac{Cov(X_1 X_2)}{\sqrt{Var(X_1)Var(X_2)}} \\
&= \frac{\frac{\sqrt{\alpha' \beta'}}{\mu'} \left({}_{\eta'}F_{\eta'}(\gamma') - \left(\frac{\gamma'}{\gamma' + 1}\right)^{\eta'} {}_{\eta'}F_{\eta'}(\gamma' + 1)\right)}{\sqrt{(\beta')^2 \left(\frac{1}{\mu'} 2 {}_{\eta'}F_{2\eta'}(\gamma') - \frac{1}{(\mu')^2} ({}_{\eta'}F_{\eta'}(\gamma'))^2\right) + \beta' \left(1 - \frac{1}{\mu'} {}_{\eta'}F_{\eta'}(\gamma')\right) + \delta'}}. \tag{39}
\end{aligned}$$

3.3 Method of moments estimation

We use the moments derived for the Lomax model in Section 3.2 to ascertain m.m.e's for the associated parameters. Nevertheless, the limited number of independent moments renders this approach inadequate for estimating $\tilde{\eta}'$. Being such, we derive moments for a $\text{Lomax}(\eta' = 1, \gamma')$ model instead.

3.3.1 Approach I

Noting the similarities in the formulations between Equation 4 in Section 2.2 and Equation 31 in Section 3.2, we may conclude that the m.m.e of α is the same as the m.m.e of α' - that is $\tilde{\alpha} = \tilde{\alpha}'$. Likewise, the similarities between Equations 5, 9 and 11 in Section 2.2 and Equations 32, 36 and 38 in Section 3.2 elucidate the equivalency of moments:

$$\tilde{\delta} + \tilde{\beta} \left(1 - \frac{1}{\tilde{\mu}} \tilde{\mu}^{\tilde{\nu}} \right) = \tilde{\delta}' + \tilde{\beta}' \left(1 - \frac{1}{\tilde{\mu}'} {}_1F_1(\tilde{\gamma}') \right), \quad (40)$$

$$\tilde{\beta}^2 \left(\frac{1}{\tilde{\mu}} \tilde{\mu}^{\tilde{\nu}^2} - \frac{1}{(\tilde{\mu})^2} \tilde{\mu}^{2\tilde{\nu}} \right) = (\tilde{\beta}')^2 \left(\frac{1}{\tilde{\mu}'} {}_2F_2(\tilde{\gamma}') - \frac{1}{(\tilde{\mu}')^2} ({}_1F_1(\tilde{\gamma}'))^2 \right), \quad (41)$$

and

$$\tilde{\beta}(1 - \tilde{\nu}) \tilde{\mu}^{\tilde{\nu}} = \tilde{\beta}' \left({}_1F_1(\tilde{\gamma}') - \left(\frac{\tilde{\gamma}'}{\tilde{\gamma}' + 1} \right) {}_1F_1(\tilde{\gamma}' + 1) \right). \quad (42)$$

Subsequently, we may relax the discreteness assumption on η' - reverting back to $\eta' > 0$ - and substitute ${}_{\eta'} F_{\eta'}(\gamma') = \sum_{i=1}^{\infty} \left(\frac{\gamma'}{\gamma'+i} \right)^{\eta'} \frac{\alpha'}{i!}$ and $\left(\frac{\gamma'}{\gamma'+1} \right)^{\eta'} {}_{\eta'} F_{\eta'}(\gamma' + 1) = \sum_{i=1}^{\infty} \left(\frac{\gamma'}{\gamma'+1+i} \right)^{\eta'} \frac{\alpha'}{i!}$ into Equations 40, 41 and 42.

Equations 40, 41 and 42 elucidate that if m.m.e's for α (hence μ), β , γ (hence ν) and δ are calculated for the exponential model using Equations 13, 16, 14 and 17, one may use Equations 40, 41 and 42 to numerically determine m.m.e's of α' , β' , γ' and δ' of the $\text{Lomax}(\eta' = 1, \gamma')$ model.

3.3.2 Approach II

Equating Equation 31 to its sample mean:

$$\begin{aligned} \tilde{\alpha}' &= M_1, \\ e^{\tilde{\alpha}'} &= e^{M_1} = \tilde{\mu}', \end{aligned} \quad (43)$$

where Equation 43 is the applicable m.m.e for all Lomax models. Consider the remaining moments, $\tilde{\beta}'$, $\tilde{\gamma}'$ and $\tilde{\delta}'$ for each model separately:

3.3.2.1 $\text{Lomax}(\eta' = 1, \gamma')$ model:

Taking the quotient of Equation 38 squared and the difference between Equation 36 and Equation 32, and equating the result to the solution obtained from identical operations on the associated sample moments, an implicit equation w.r.t $\tilde{\gamma}'$ is obtained, to be

solved numerically:

$$\frac{(\tilde{\alpha}')^2 \left[{}_1F_1(\tilde{\gamma}') - \left(\frac{\tilde{\gamma}'}{\tilde{\gamma}' + 1} \right) {}_1F_1(\tilde{\gamma}' + 1) \right]^2}{\tilde{\mu}' {}_2F_2(\tilde{\gamma}') - [{}_1F_1(\tilde{\gamma}')]^2} = \frac{{S_{12}}^2}{S_2 - M_2}, \quad (44)$$

and to ensure the expression remains real-valued we obtain the condition:

$$S_2 \neq M_2.$$

Since $\forall x_1, x_2 \in 0, 1, 2, \dots$ and $\tilde{\gamma}' > 0$, it is necessary to establish bounds for the quantity $\frac{{S_{12}}^2}{S_2 - M_2}$ to ensure that the condition $\tilde{\gamma}' > 0$ is satisfied.

Now, since the left-hand side of Equation 44 is strictly increasing in $\tilde{\gamma}' > 0$ (as proven in Theorem 4), we may take the limits $\lim_{\tilde{\gamma}' \rightarrow 0^+}$ and $\lim_{\tilde{\gamma}' \rightarrow \infty}$ of the left-hand side of Equation 44 to obtain its lower and upper bounds, respectively. Furthermore, by substituting using Equations 40, 41, and 42 into the left-hand side of Equation 44, we recover the left-hand side of Equation 14. Since we already established that taking the limits $\lim_{\tilde{\gamma} \rightarrow \infty}$ and $\lim_{\tilde{\gamma} \rightarrow 0^+}$ in the exponential model is equivalent to taking $\lim_{\tilde{\gamma}' \rightarrow 0^+}$ and $\lim_{\tilde{\gamma}' \rightarrow \infty}$ in the Lomax model, it is evident that we obtain the same lower and upper bounds for the left-hand side of Equation 14, as given in Equation 15 (we show $\lim_{\tilde{\gamma}' \rightarrow 0^+}$ of the left-side expression in Equation 44 in in Theorem 5 regardless). Thus, the bounds for the quantity $\frac{{S_{12}}^2}{S_2 - M_2}$ are as follows (equivalent to Equation 15):

$$\frac{M_1^2}{e^{M_1} - 1} < \frac{{S_{12}}^2}{S_2 - M_2} < M_1. \quad (45)$$

Figure 5 illustrates $\tilde{\gamma}'$ as a function of $\tilde{\alpha} = M_1$ and $\frac{{S_{12}}^2}{S_2 - M_2}$. Importantly, due to the bounds imposed by M_1 on $\frac{{S_{12}}^2}{S_2 - M_2}$ in Equation 45, certain combinations of M_1 , S_2 , M_2 and S_{12} values may lead to the non-existence of $\tilde{\gamma}'$. This behavior is also evident from the patterns observed in Figure 5.

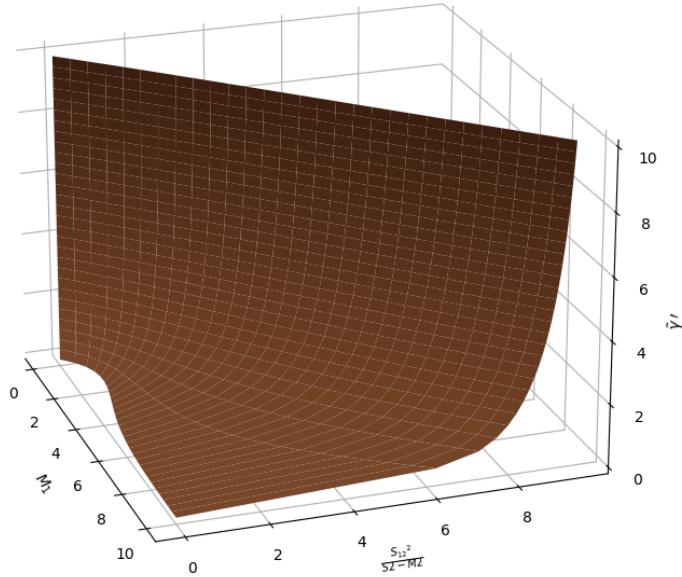


Figure 5: $\tilde{\gamma}'$ vs M_1 and $\frac{{S_{12}}^2}{S_2 - M_2}$ as per Equation 44.

We then solve for $\tilde{\beta}'$ by equating Equation 38 to the sample covariance to obtain:

$$\tilde{\beta}' = \frac{S_{12}}{\frac{\tilde{\alpha}'}{\tilde{\mu}'} \left({}_1F_1(\tilde{\gamma}') - \left(\frac{\tilde{\gamma}'}{\tilde{\gamma}' + 1} \right) {}_1F_1(\tilde{\gamma}' + 1) \right)}.$$

Finally, solve for $\tilde{\delta}'$ by merely equating Equation 32 to the sample mean of X_2 :

$$\tilde{\delta}' = M_2 - \tilde{\beta}' \left(1 - \frac{1}{\tilde{\mu}'} {}_1F_1(\tilde{\gamma}') \right).$$

We note that the m.m.e's of Case 3: $\delta' = 0$, may be derived analogously to the approach previously outlined, with the natural exception of excluding the estimator for δ' .

3.3.2.2 Case I ($\beta' = 1$ or $\beta' = -1$):

With regards to solving for $\tilde{\gamma}'$, one may either solve Equation 44, or alternatively, equate the model-based expression for $Cov(X_1, X_2)$ (Equation 11) with the sample covariance. This yields:

$$\text{sign}(\beta') \frac{\tilde{\alpha}'}{\tilde{\mu}'} \left({}_1F_1(\tilde{\gamma}') - \left(\frac{\tilde{\gamma}'}{\tilde{\gamma}' + 1} \right) {}_1F_1(\tilde{\gamma}' + 1) \right) = S_{12}, \quad (46)$$

or equivalently:

$$\text{sign}(\beta') \frac{\tilde{\alpha}'}{\tilde{\mu}'} \sum_{i=0}^{\infty} \frac{\tilde{\gamma}'}{(\tilde{\gamma}' + i)(\tilde{\gamma}' + 1 + i)} \frac{(\tilde{\alpha}')^i}{i!} = S_{12}.$$

Since $\forall x_1, x_2 \in \{0, 1, 2, \dots\}$, $\tilde{\gamma}' > 0$, it follows that bounds must be established for the quantity S_{12} in order to guarantee the condition $\tilde{\gamma}' > 0$ is satisfied. Since the left-hand side expression of Equation 46 is not monotonic with respect to $\tilde{\gamma}'$, it is not possible to determine bounds for S_{12} by using limits to guarantee $\tilde{\gamma}' > 0$, as was done previously. Figure 6 illustrates the relationship between $\tilde{\gamma}'$, $\tilde{\alpha}' = M_1$, and $\text{sign}(\beta')S_{12}$. Importantly, due to the inability to establish definitive bounds on $\text{sign}(\beta')S_{12}$ ensuring a positive m.m.e. for γ' , certain combinations of M_1 and S_{12} values may lead to the non-existence of $\tilde{\gamma}'$. This behavior is also evident from the patterns observed in Figure 6.

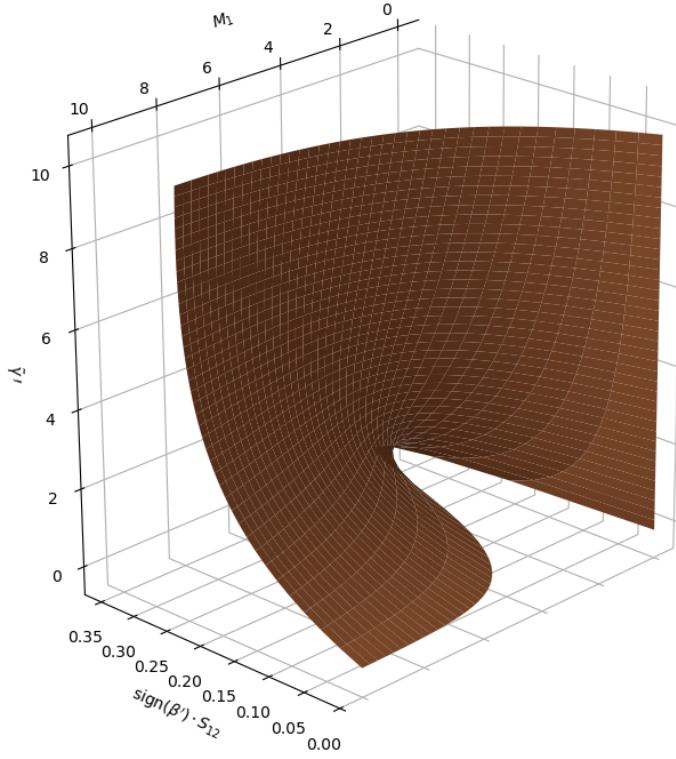


Figure 6: $\tilde{\gamma}'$ vs M_1 and $\text{sign}(\beta')$ as per Equation 46.

Furthermore, we solve for $\tilde{\delta}'$ just as before, by merely equating Equation 32 to the sample mean of X_2 :

$$\tilde{\delta}' = M_2 - \text{sign}(\beta') \left(1 - \frac{1}{\tilde{\mu}'} {}_1 F_1(\tilde{\gamma}') \right).$$

3.3.2.3 Case II ($\gamma' = 1$):

Since $\gamma' = 1$, we have ${}_1 F_1(\tilde{\gamma}' = 1) = \frac{1}{\tilde{\alpha}'} (\tilde{\mu}' - 1)$ and ${}_1 F_1(\tilde{\gamma} + 1 = 2) = \frac{1}{(\tilde{\alpha}')^2} (\tilde{\mu}'(\tilde{\alpha}' - 1) + 1)$. Now we equate the model-based expression for $Cov(X_1, X_2)$ (Equation 38) with the sample covariance to obtain:

$$\tilde{\beta}' = \frac{S_{12}}{\frac{1}{\tilde{\alpha}' \tilde{\mu}'} (\tilde{\mu}' - \tilde{\alpha}' - 1)}. \quad (47)$$

Afterwhich, we solve for $\tilde{\delta}'$ equating Equation 32 to the sample mean of X_2 :

$$\tilde{\delta}' = M_2 - \tilde{\beta}' \left(1 - \frac{1}{\tilde{\alpha}'} \left(1 - \frac{1}{\tilde{\mu}'} \right) \right).$$

3.3.2.4 Case IV ($\gamma' = 1$ and $\beta' = 1$ or $\gamma' = 1$ and $\beta' = -1$):

We solve for $\tilde{\delta}'$ by merely equating Equation 32 to the sample mean of X_2 :

$$\tilde{\delta}' = M_2 - \text{sign}(\beta') \left(1 - \frac{1}{\tilde{\alpha}'} \left(1 - \frac{1}{\tilde{\mu}'} \right) \right).$$

3.3.2.5 Case V ($\delta' = 0$ and $\gamma' = 1$):

We solve for $\tilde{\beta}'$ by merely equating Equation 32 to the sample mean of X_2 :

$$\tilde{\beta}' = \frac{M_2}{1 - \frac{1}{\mu'} {}_1F_1(\tilde{\gamma}')}. \quad (39)$$

3.4 Correlations

We note, as previously explained, the equivalence of the Lomax model with the exponential model, specifically that $\alpha = \alpha'$, $\beta = \beta'$, and $\delta = \delta'$ under the limit $\lim_{\gamma' \rightarrow 0^+}$, which corresponds to taking the limit $\lim_{\gamma \rightarrow \infty}$ in the exponential model. Consequently, the correlation bounds, for the full Lomax model, the Lomax($\eta' = 1, \gamma'$) model and its sub-models, Case I and Case III, coincide with those of their exponential model counterparts. Referring to Equations 18 and 19 for the full exponential model, and Equations 20 and 21 for Case I, we observe that the limit $\lim_{\gamma \rightarrow \infty}$ applied to the correlation function of the full exponential model in Equation 12 directly establishes the equivalence between the exponential and Lomax models. However, this equivalence does not extend to Cases II, IV and V of the exponential model, and therefore, the correlation bounds for the Lomax sub-models Case II, Case IV and Case V require separate derivation. Furthermore, we utilise the Lomax correlation function $\rho(\alpha', \beta', \gamma', \delta', \eta' = 1)$ given in Equation 39 in this section.

3.5 Case II

Since $\gamma' = 1$, we have ${}_1F_1(\gamma' = 1) = \frac{1}{\alpha'}(\mu' - 1)$ and ${}_1F_1(\gamma' + 1 = 2) = \frac{1}{(\alpha')^2}(\mu'(\alpha' - 1) + 1)$. Now consider:

$$\begin{aligned} {}_2F_2(\gamma' = 1) &= \sum_{i=0}^{\infty} \left(\frac{1}{1+i} \right)^2 \frac{\alpha'}{i!} \\ &= \frac{1}{\alpha'} \sum_{j=1}^{\infty} \frac{1}{j} \frac{(\alpha')^j}{j!} \\ &= \frac{1}{\alpha'} (\text{Ei}(\alpha') - \gamma^* - \log|\alpha'|), \end{aligned}$$

where $\text{Ei}(\alpha')$ is the exponential integral with γ^* being the Euler-Mascheroni constant. Hence from Equation 39,

$$\rho(\alpha', \beta', \gamma' = 1, \delta', \eta' = 1) = \frac{\beta' \left(\mu' - 1 - \frac{(\alpha' - 1)\mu'}{\alpha'} - \frac{1}{\alpha'} \right)}{\sqrt{\alpha' \left(\beta' \mu' \left(\mu' - \frac{\mu'}{\alpha} + \frac{1}{\alpha} \right) + \delta' (\mu')^2 + \frac{(\beta')^2 \left(\mu'(-\log(\alpha') + \text{Ei}(\alpha') - \gamma^*) - \frac{(1-\mu')^2}{\alpha'} \right)}{\alpha'} \right)}}. \quad (40)$$

Now after taking the limit:

$$\lim_{\beta' \rightarrow \infty} \rho(\alpha', \beta', \gamma' = 1, \delta', \eta' = 1) = \frac{\sqrt{\alpha'} \left(\mu' - 1 - \frac{(\alpha' - 1)\mu'}{\alpha'} - \frac{1}{\alpha'} \right)}{\sqrt{A}}, \quad (41)$$

and

$$\lim_{\beta' \rightarrow -\infty} \rho(\alpha', \beta', \gamma' = 1, \delta', \eta' = 1) = -\frac{\sqrt{\alpha'} \left(\mu' - 1 - \frac{(\alpha' - 1)\mu'}{\alpha'} - \frac{1}{\alpha'} \right)}{\sqrt{A}}, \quad (42)$$

with $A = -\alpha'\mu' \log |\alpha'| + \alpha'\mu' \text{Ei}(\alpha') - \gamma^* \alpha' \mu' - (\mu')^2 + 2\mu' - 1$. After which, taking $\lim_{\alpha' \rightarrow 0^+}$ to Equations 48 and 49 implies $-1 < \rho < 1$.

3.6 Case IV

3.6.0.1 $\beta' = 1$

Again, taking the limit of Equation 39, and noting for positive correlation (that is, $\beta' > 0$) we must have $\delta' \geq 0$:

$$\lim_{\delta' \rightarrow 0^+} \rho(\alpha', \beta' = 1, \gamma' = 1, \delta', \eta' = 1) = \frac{\mu' - 1 + \frac{(1-\alpha)\mu'}{\alpha} - \frac{1}{\alpha}}{\sqrt{A + (\alpha'\mu')^2 - \alpha'(\mu')^2 + \alpha'\mu'}}. \quad (50)$$

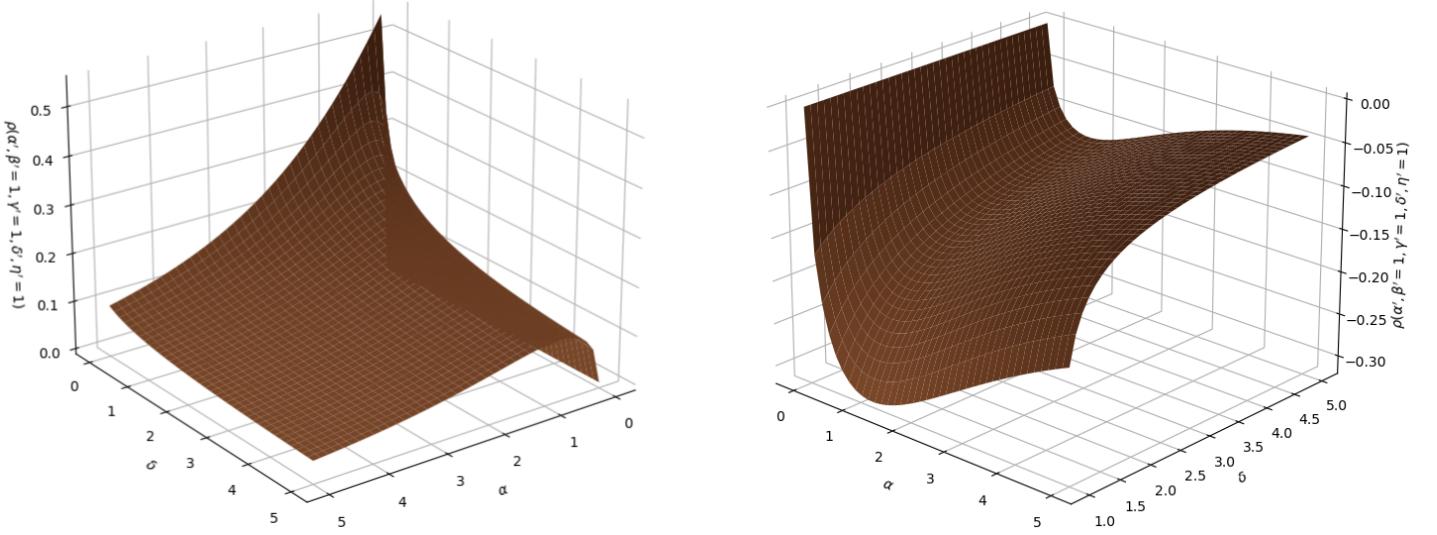
Noting the monotonicity of Equation 50, we take $\lim_{\alpha' \rightarrow 0^+}$ to obtain $\frac{\sqrt{3}}{3} \approx 0.57735$. Hence, our bounds are $0 < \rho < \frac{\sqrt{3}}{3} \approx 0.57735$ also illustrated in Figure 7a displaying $\rho(\alpha', \beta' = 1, \gamma' = 1, \delta', \eta' = 1)$.

3.6.0.2 $\beta = -1$

Again, taking the limit of Equation 39, and noting for negative correlation (that is, $\beta' < 0$) we must have $\delta' \geq -\beta$:

$$\lim_{\delta' \rightarrow 1^+} \rho(\alpha', \beta' = 1, \gamma' = 1, \delta', \eta' = 1) = -\frac{\sqrt{\alpha'} \left(\mu' - 1 - \frac{(1-\alpha)\mu'}{\alpha} - \frac{1}{\alpha} \right)}{\sqrt{A + \alpha'(\mu')^2 - \alpha'\mu'}} \quad (51)$$

Now Equation 51 is not monotonic with respect to α , so instead of taking limits, we rather minimise computationally: $(\alpha')^* = \operatorname{argmin}_{\alpha'} \{\lim_{\delta' \rightarrow 1} \rho(\alpha', \beta' = 1, \gamma' = 1, \delta', \eta' = 1)\} = 1.2345$, such that $\{\lim_{\delta' \rightarrow 1} \rho(\alpha', \beta' = 1, \gamma' = 1, \delta', \eta' = 1)\}_{|\alpha'=(\alpha')^*} \approx -0.31465$. Hence, our bounds are $\approx -0.31465 < \rho < 0$ also illustrated in Figure 7b displaying $\rho(\alpha', \beta' = -1, \gamma' = 1, \delta', \eta' = 1)$.



(a) $\rho(\alpha', \beta' = 1, \gamma' = 1, \delta', \eta' = 1)$ against $\alpha' > 0$ and $\delta' \geq 0$ noting a maximum achieved at $\frac{\sqrt{3}}{3}$.

(b) $\rho(\alpha', \beta' = -1, \gamma' = 1, \delta', \eta' = 1)$ against $\alpha' > 0$, $\delta' \geq 1$ noting a minimum achieved at ≈ -0.314 .

Figure 7: $\rho(\alpha', \beta' = 1, \gamma' = 1, \delta', \eta' = 1)$ and $\rho(\alpha', \beta' = -1, \gamma' = 1, \delta', \eta' = 1)$.

3.7 Case V

Similarly, taking the limits of Equation 39:

$$\lim_{\beta' \rightarrow \infty} \lim_{\alpha' \rightarrow 0^+} \rho(\alpha', \beta', \gamma' = 1, \delta' = 0) = 1,$$

and

$$\lim_{\beta' \rightarrow -\infty} \lim_{\alpha' \rightarrow 0^+} \rho(\alpha', \beta', \gamma' = 1, \delta' = 0) = -1.$$

Hence, the bounds are $-1 < \rho < 1$.

We summarise the correlation $\rho(\alpha', \beta', \gamma', \delta', \eta')$ of Equation 39 bounds in Table 2, confirmed with numerical optimisation.

Model	ρ	Lower Bound	Upper Bound
Full Lomax, $\text{Lomax}(\eta' = 1, \gamma')$ (including Case III)		-1	1
Case I ($\beta' = 1$)		0	$\frac{\sqrt{2}}{2} \approx 0.707$
Case I ($\beta' = -1$)		≈ -0.482	0
Case II ($\gamma' = 1$)		-1	1
Case IV ($\beta' = 1, \gamma' = 1$)		0	$\frac{\sqrt{3}}{3} \approx 0.577$
Case IV ($\beta' = -1, \gamma' = 1$)		≈ -0.315	0
Case V ($\delta' = 0, \gamma' = 1$)		-1	1

Table 2: ρ lower and upper bounds for the full lomax model and sub-models.

3.8 Maximum likelihood estimation

Consider again, the data in the form of $\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \dots, \mathbf{X}^{(n)}$ (where for each $i = 1, \dots, n$, $\mathbf{X}^{(i)} = (X_{1i}, X_{2i})^T$) are obtained, which are independent and identically distributed according to Equations 28 - 29, the the likelihood function is as follows from Equation 30:

$$L(\alpha', \beta', \gamma', \delta', \eta') = \prod_{i=1}^n \frac{e^{-\alpha'(x_{1i})} e^{-\delta' - \beta' \left(1 - \left(\frac{\gamma'}{x_{1i} + \gamma'}\right)^{\eta'}\right)}}{x_{1i}!} \frac{\left\{ \delta' + \beta' \left(1 - \left(\frac{\gamma'}{x_{1i} + \gamma'}\right)^{\eta'}\right) \right\}^{x_{2i}}}{x_{2i}!}$$

$$= \frac{e^{-n(\alpha' + \beta' + \delta')} (\alpha')^{\sum_{i=1}^n x_{1i}} e^{\beta' \sum_{i=1}^n \left(\frac{\gamma'}{x_{1i} + \gamma'}\right)^{\eta'}} \prod_{i=1}^n \left(\delta' + \beta' \left(1 - \left(\frac{\gamma'}{x_{1i} + \gamma'}\right)^{\eta'}\right) \right)^{x_{2i}}}{\prod_{i=1}^n (x_{1i}! x_{2i}!)}$$

Now the log-likelihood is:

$$l = \log(L(\alpha', \beta', \gamma', \delta', \eta')) = -n(\alpha' + \beta' + \delta') + \log(\alpha') \sum_{i=1}^n x_{1i} + \beta' \sum_{i=1}^n \left(\frac{\gamma'}{x_{1i} + \gamma'}\right)^{\eta'} + \sum_{i=1}^n x_{2i} \log \left(\delta' + \beta' \left(1 - \left(\frac{\gamma'}{x_{1i} + \gamma'}\right)^{\eta'}\right) \right) + h(\mathbf{x}_1, \mathbf{x}_2), \quad (52)$$

where $h(\mathbf{x}_1, \mathbf{x}_2) = -\log(\prod_{i=1}^n (x_{1i}! x_{2i}!))$. Now since $\frac{\partial l}{\partial \alpha'} = 0 \implies \hat{\alpha}' = \frac{\sum_{i=1}^n x_{1i}}{n} = \tilde{\alpha}'$, that is, the m.l.e of α' is equivalent to the m.m.e of α' . The m.l.e's of $\beta', \gamma', \delta', \eta'$ are to be solved numerically using Equation 52

3.9 Likelihood ratio tests

We note the natural parameter space of the full Lomax model is $\Theta = \{(\alpha', \beta', \gamma', \delta', \eta')^T : \alpha' > 0, \gamma' > 0, \beta' \in \mathbb{R} \setminus \{0\}, \delta' \geq \max(-\beta', 0), \eta' > 0\}$. Let $\hat{\alpha}', \hat{\beta}', \hat{\gamma}', \hat{\delta}'$ and $\hat{\eta}'$ be the m.l.e's of the full Lomax model. Furthermore, we denote $(\hat{\alpha}')^*, (\hat{\beta}')^*, (\hat{\gamma}')^*, (\hat{\delta}')^*$ and $(\hat{\eta}')^*$ as any m.l.e under H_0 in this section.

3.9.1 Testing $\beta' = \pm 1$

Under H_0 , the natural parameter space is $\Theta_0 = \{(\alpha', \gamma', \delta', \eta')^T : \alpha' > 0, \gamma' > 0, \delta' \geq \mathbb{I}(\beta' < 0), \eta' > 0\}$. The generalized likelihood ratio test statistic in Equation 27 is:

$$\Lambda = \frac{\frac{e^{-n((\hat{\alpha}')^* + \text{sign}(\beta') + (\hat{\delta}')^*)} ((\hat{\alpha}')^*)^{\sum_{i=1}^n x_{1i}} e^{\text{sign}(\beta') \sum_{i=1}^n \left(\frac{(\hat{\gamma}')^*}{x_{1i} + (\hat{\gamma}')^*}\right)^{(\hat{\eta}')^*}} \prod_{i=1}^n \left((\hat{\delta}')^* + (\hat{\beta}')^* \left(1 - \left(\frac{(\hat{\gamma}')^*}{x_{1i} + (\hat{\gamma}')^*}\right)^{(\hat{\eta}')^*}\right) \right)^{x_{2i}}}{\prod_{i=1}^n (x_{1i}! x_{2i}!)}}{\frac{e^{-n(\hat{\alpha}' + \hat{\beta}' + \hat{\delta}')(\hat{\alpha}')^*} (\hat{\alpha}')^{\sum_{i=1}^n x_{1i}} e^{\hat{\beta}' \sum_{i=1}^n \left(\frac{(\hat{\gamma}')^*}{x_{1i} + \hat{\gamma}'}\right)^{\hat{\eta}'}} \prod_{i=1}^n \left(\hat{\delta}' + \hat{\beta}' \left(1 - \left(\frac{(\hat{\gamma}')^*}{x_{1i} + \hat{\gamma}'}\right)^{\hat{\eta}'}\right) \right)^{x_{2i}}}{\prod_{i=1}^n (x_{1i}! x_{2i}!)}}.$$

Since $(\hat{\alpha}')^* = \hat{\alpha}'$, we have:

$$\Lambda = e^{-n(\text{sign}(\beta') + (\hat{\delta}')^* - \hat{\beta}' - \hat{\delta}')} e^{\left(\text{sign}(\beta') \sum_{i=1}^n \left(\frac{(\hat{\gamma}')^*}{x_{1i} + (\hat{\gamma}')^*}\right)^{(\hat{\eta}')^*} - \hat{\beta}' \sum_{i=1}^n \left(\frac{(\hat{\gamma}')^*}{x_{1i} + (\hat{\gamma}')^*}\right)^{\hat{\eta}'}\right)} \prod_{i=1}^n \left[\frac{(\hat{\delta}')^* + \text{sign}(\beta') \left(1 - \left(\frac{(\hat{\gamma}')^*}{x_{1i} + (\hat{\gamma}')^*}\right)^{(\hat{\eta}')^*}\right)}{\hat{\delta}' + \hat{\beta}' \left(1 - \left(\frac{(\hat{\gamma}')^*}{x_{1i} + \hat{\gamma}'}\right)^{\hat{\eta}'}\right)} \right]^{x_{2i}}.$$

And by taking the logarithm, we obtain:

$$\begin{aligned} \log \Lambda = & -n \left(\text{sign}(\beta') + (\hat{\delta}')^* - \hat{\beta}' - \hat{\delta}' \right) + \text{sign}(\beta') \sum_{i=1}^n \left(\frac{(\hat{\gamma}')^*}{x_{1i} + (\hat{\gamma}')^*} \right)^{(\hat{\eta}')^*} - \hat{\beta}' \sum_{i=1}^n \left(\frac{(\hat{\gamma}')^*}{x_{1i} + (\hat{\gamma}')^*} \right)^{\hat{\eta}'} \\ & + \sum_{i=1}^n x_{2i} \log \left[\frac{(\hat{\delta}')^* + \text{sign}(\beta') \left(1 - \left(\frac{(\hat{\gamma}')^*}{x_{1i} + (\hat{\gamma}')^*}\right)^{(\hat{\eta}')^*}\right)}{\hat{\delta}' + \hat{\beta}' \left(1 - \left(\frac{(\hat{\gamma}')^*}{x_{1i} + \hat{\gamma}'}\right)^{\hat{\eta}'}\right)} \right]. \end{aligned}$$

3.9.2 Testing $\gamma' = 1$

Under H_0 , the natural parameter space is $\Theta_0 = \{(\alpha', \beta', \delta', \eta')^T : \alpha' > 0, \beta' \in \mathbb{R} \setminus \{0\}, \delta' \geq \max(-\beta, 0), \eta' > 0\}$. The logarithm of the generalized likelihood ratio test statistic in Equation 27 is:

$$\begin{aligned} \log \Lambda = & -n \left((\hat{\beta}')^* + (\hat{\delta}')^* - \hat{\beta}' - \hat{\delta}' \right) + (\hat{\beta}')^* \sum_{i=1}^n \left(\frac{1}{x_{1i} + 1} \right)^{(\hat{\eta}')^*} - \hat{\beta}' \sum_{i=1}^n \left(\frac{(\hat{\gamma}')^*}{x_{1i} + (\hat{\gamma}')^*} \right)^{\hat{\eta}'} \\ & + \sum_{i=1}^n x_{2i} \log \left[\frac{(\hat{\delta}')^* + (\hat{\beta}')^* \left(1 - \left(\frac{1}{x_{1i} + 1}\right)^{(\hat{\eta}')^*}\right)}{\hat{\delta}' + \hat{\beta}' \left(1 - \left(\frac{(\hat{\gamma}')^*}{x_{1i} + \hat{\gamma}'}\right)^{\hat{\eta}'}\right)} \right]. \end{aligned}$$

3.9.3 Testing $\delta' = 0$

Under H_0 , the natural parameter space is $\Theta_0 = \{(\alpha', \beta', \gamma', \eta')^T : \alpha' > 0, \beta' \in \mathbb{R} \setminus \{0\}, \gamma' > 0, \eta' > 0\}$. The logarithm of the generalized likelihood ratio test statistic in Equation 27 is:

$$\begin{aligned} \log \Lambda = & -n \left((\hat{\beta}')^* - \hat{\beta}' - \hat{\delta}' \right) + (\hat{\beta}')^* \sum_{i=1}^n \left(\frac{(\hat{\gamma}')^*}{x_{1i} + (\hat{\gamma}')^*} \right)^{(\hat{\eta}')^*} - \hat{\beta}' \sum_{i=1}^n \left(\frac{(\hat{\gamma}')^*}{x_{1i} + (\hat{\gamma}')^*} \right)^{\hat{\eta}'} \\ & + \sum_{i=1}^n x_{2i} \log \left[\frac{(\hat{\beta}')^* \left(1 - \left(\frac{(\hat{\gamma}')^*}{x_{1i} + (\hat{\gamma}')^*}\right)^{(\hat{\eta}')^*}\right)}{\hat{\delta}' + \hat{\beta}' \left(1 - \left(\frac{(\hat{\gamma}')^*}{x_{1i} + \hat{\gamma}'}\right)^{\hat{\eta}'}\right)} \right]. \end{aligned}$$

3.9.4 Testing $\beta' = \pm 1$ and $\gamma' = 1$

Under H_0 , the natural parameter space is $\Theta_0 = \{(\alpha', \delta', \eta')^T : \alpha' > 0, \delta' \geq \mathbb{I}(\beta' < 0), \eta' > 0\}$. The logarithm of the generalized likelihood ratio test statistic in Equation 27 is:

$$\begin{aligned}\log \Lambda = & -n \left(\text{sign}(\beta') + (\hat{\delta}')^* - \hat{\beta}' - \hat{\delta}' \right) + \text{sign}(\beta') \sum_{i=1}^n \left(\frac{1}{x_{1i} + 1} \right)^{(\hat{\eta}')^*} - \hat{\beta}' \sum_{i=1}^n \left(\frac{(\hat{\gamma}')}{x_{1i} + (\hat{\gamma}')} \right)^{\hat{\eta}'} \\ & + \sum_{i=1}^n x_{2i} \log \left[\frac{(\hat{\delta}')^* + \text{sign}(\beta') \left(1 - \left(\frac{1}{x_{1i} + 1} \right)^{(\hat{\eta}')^*} \right)}{\hat{\delta}' + \hat{\beta}' \left(1 - \left(\frac{\hat{\gamma}'}{x_{1i} + \hat{\gamma}'} \right)^{\hat{\eta}'} \right)} \right].\end{aligned}$$

3.9.5 Testing $\delta' = 0$ and $\gamma' = 1$

Under H_0 , the natural parameter space is $\Theta_0 = \{(\alpha', \beta', \eta')^T : \alpha' > 0, \beta' \in \mathbb{R} \setminus \{0\}, \eta' > 0\}$. The logarithm of the generalized likelihood ratio test statistic in Equation 27 is:

$$\begin{aligned}\log \Lambda = & -n \left((\hat{\beta}')^* - \hat{\beta}' - \hat{\delta}' \right) + (\hat{\beta}')^* \sum_{i=1}^n \left(\frac{1}{x_{1i} + 1} \right)^{(\hat{\eta}')^*} - \hat{\beta}' \sum_{i=1}^n \left(\frac{(\hat{\gamma}')}{x_{1i} + (\hat{\gamma}')} \right)^{\hat{\eta}'} \\ & + \sum_{i=1}^n x_{2i} \log \left[\frac{(\hat{\beta}')^* \left(1 - \left(\frac{1}{x_{1i} + 1} \right)^{(\hat{\eta}')^*} \right)}{\hat{\delta}' + \hat{\beta}' \left(1 - \left(\frac{\hat{\gamma}'}{x_{1i} + \hat{\gamma}'} \right)^{\hat{\eta}'} \right)} \right].\end{aligned}$$

3.9.6 Testing $\eta' = 1$

Under H_0 , the natural parameter space is $\Theta_0 = \{(\alpha', \beta', \gamma', \delta')^T : \alpha' > 0, \beta' \in \mathbb{R} \setminus \{0\}, \gamma' > 0, \delta' \geq \max(-\beta, 0)\}$. The logarithm of the generalized likelihood ratio test statistic in Equation 27 is:

$$\begin{aligned}\log \Lambda = & -n \left((\hat{\beta}')^* + (\hat{\delta}')^* - \hat{\beta}' - \hat{\delta}' \right) + (\hat{\beta}')^* \sum_{i=1}^n \frac{(\hat{\gamma}')^*}{x_{1i} + (\hat{\gamma}')^*} - \hat{\beta}' \sum_{i=1}^n \left(\frac{(\hat{\gamma}')}{x_{1i} + (\hat{\gamma}')} \right)^{\hat{\eta}'} \\ & + \sum_{i=1}^n x_{2i} \log \left[\frac{(\hat{\delta}')^* + (\hat{\beta}')^* \left(1 - \frac{(\hat{\gamma}')^*}{x_{1i} + (\hat{\gamma}')^*} \right)}{\hat{\delta}' + \hat{\beta}' \left(1 - \left(\frac{\hat{\gamma}'}{x_{1i} + \hat{\gamma}'} \right)^{\hat{\eta}'} \right)} \right].\end{aligned}$$

4 Simulation Data

Simulating from the pseudo-models is straightforward because of their marginal-conditional structure. For example, for the full exponential model given $(\alpha, \beta, \gamma, \delta)$, we generate a bivariate sample (x_1, x_2) as follows:

Step I: Simulate x_1 from $\text{Poisson}(\alpha)$.

Step II: Simulate x_2 from $\text{Poisson}(\delta + \beta(1 - e^{\gamma x_1}))$.

We also simulate data for the exponential Case I ($\rho < 0$) sub-model by fixing $\beta = -1$. An analogous simulation procedure applies to the Lomax model and its sub-models. Furthermore, we simulate 10000 data sets of size $n = 100, 1000$ and 10000 for fixed α, β, γ and δ values.

For the full exponential model (Table 3), recall that the bounds implied by M_1 on $\frac{S_{12}^2}{S_2 - M_2}$ in Equation 15 may lead to the

non-existence of the m.m.e. $\tilde{\gamma}$ for certain combinations of M_1 , S_2 , M_2 , and S_{12} (see Figure 2). To avoid this issue, we select parameter values that satisfy these bounds, specifically: $\alpha = 5$, $\beta = -20$, $\gamma = 0.5$ and $\delta = 25$. An analogous construction applies to the Lomax($\eta' = 1, \gamma'$) model (Table 5), where existence of $\tilde{\gamma}'$ is governed by the bounds in Equation 45. Similarly, recall for the exponential Case I ($\rho < 0$) sub-model (Table 4), the existence of a real-valued solution for $\tilde{\nu}$ requires that $\frac{M_1}{S_{12}}\text{sign}(\beta) \geq e$ - we therefore select parameter values to ensure this condition is met, setting them as: $\alpha = 1$, $\gamma = 0.3$ and $\delta = 25$.

We omit simulation results for the remaining sub-models, as no bounds are required to guarantee the existence of their m.m.e's; consequently, their corresponding simulations are straightforward. Furthermore, recall that no formal existence bounds for $\tilde{\gamma}'$ could be derived for the Lomax Case I sub-model; instead, Figure 6 was provided solely to facilitate a visual illustration of the existence of $\tilde{\gamma}'$. Accordingly, simulations for the Lomax Case I sub-model are not reported here - in practical applications, the existence of $\tilde{\gamma}'$ would depend exclusively on the convergence of the optimisation routine employed.

The corresponding m.m.e and m.l.e results are reported in Tables 3 and 4 for the full exponential model and exponential Case I sub-model, and Table 5 for the Lomax($\eta' = 1, \gamma'$) model. Included are 95% confidence intervals computed as $\hat{\theta} \pm Z_{\alpha/2}\text{S.E.}(\hat{\theta})^1$, where $\hat{\theta}$ is the point estimator (as done in Arnold and Manjunath [1]). As anticipated, the standard errors of both the m.m.e. and m.l.e. decrease with increasing sample size n , resulting in correspondingly narrower 95% confidence intervals. For the full exponential model (Table 3) and the Lomax($\eta' = 1, \gamma'$) model (Table 5), the m.l.e.'s consistently outperform the m.m.e.'s, yielding smaller standard errors and tighter confidence intervals. In contrast, for the exponential Case I sub-model ($\rho < 0$), the reverse pattern is observed, with the m.m.e.'s demonstrating greater accuracy. A salient feature in Table 4 is the relative instability of the estimators of γ : reliable estimation of γ appears to require substantially larger sample sizes. Finally, we note that the Pearson correlation (PC) converges to the population correlation as n increases.

n	Parameter	Moment	MLE	SE(Moment)	SE(MLE)	95%CI (MM)	95%CI (MM)	PC
100	α	5.000	5.000	0.217	0.217	(4.574, 5.426)	(4.574, 5.426)	
	β	-19.493	-20.269	4.533	3.512	(-28.378, -10.607)	(-27.153, -13.385)	
	γ	0.474	0.504	0.164	0.102	(0.154, 0.795)	(0.304, 0.703)	-0.601
	δ	23.929	25.160	5.078	3.827	(13.977, 33.882)	(17.660, 32.661)	
	ρ	-0.596	-0.597	0.075	0.044	(-0.743, -0.449)	(-0.683, -0.511)	
1000	α	5.000	5.000	0.069	0.069	(4.864, 5.134)	(4.864, 5.134)	
	β	-19.872	-20.010	1.638	1.082	(-23.083, -16.660)	(-22.131, -17.889)	
	γ	0.496	0.500	0.054	0.032	(0.391, 0.601)	(0.437, 0.563)	-0.599
	δ	24.853	25.002	1.768	1.165	(21.388, 28.319)	(22.718, 27.285)	
	ρ	-0.598	-0.598	0.024	0.014	(-0.646, -0.550)	(-0.625, -0.571)	
10000	α	5.000	5.000	0.022	0.022	(4.957, 5.042)	(4.957, 5.042)	
	β	-19.996	-20.002	0.523	0.348	(-21.022, -18.970)	(-20.683, -19.320)	
	γ	0.500	0.500	0.017	0.010	(0.466, 0.533)	(0.480, 0.520)	-0.598
	δ	24.990	25.003	0.563	0.375	(23.885, 26.094)	(24.268, 25.739)	
	ρ	-0.598	-0.598	0.008	0.004	(-0.613, -0.583)	(-0.606, -0.589)	

Table 3: Simulations for full exponential model ($\alpha = 5$, $\beta = -20$, $\gamma = 0.5$, $\delta = 25$).

¹ $Z_{\alpha/2}$ denotes the $100(1 - \alpha/2)^{\text{th}}$ percentile of the standard normal distribution.

n	Parameter	Moment	MLE	SE(Moment)	SE(MLE)	95%CI (MM)	95%CI (MM)	PC
100	α	1.000	1.000	0.096	0.096	(0.812, 1.188)	(0.812, 1.188)	-0.038
	γ	0.247	0.196	0.273	1.506	(-0.288, 0.782)	(-2.755, 3.147)	
	δ	24.997	25.054	0.504	0.544	(24.010, 25.985)	(23.989, 26.119)	
	ρ	-0.037	-0.040	0.021	0.027	(-0.078, 0.004)	(-0.093, 0.012)	
1000	α	1.000	1.000	0.031	0.031	(0.940, 1.060)	(0.940, 1.060)	-0.038
	γ	0.265	0.214	0.240	0.282	(-0.206, 0.736)	(-0.338, 0.767)	
	δ	24.997	24.988	0.206	0.257	(24.593, 25.402)	(24.485, 25.492)	
	ρ	-0.039	-0.037	0.019	0.023	(-0.077, -0.001)	(-0.083, 0.009)	
10000	α	1.000	1.000	0.010	0.010	(0.981, 1.019)	(0.981, 1.019)	-0.040
	γ	0.297	0.251	0.113	0.195	(0.076, 0.518)	(-0.131, 0.633)	
	δ	25.002	24.975	0.082	0.154	(24.842, 25.163)	(24.672, 25.277)	
	ρ	-0.040	-0.037	0.010	0.019	(-0.059, -0.021)	(-0.074, -0.001)	

Table 4: Simulations for exponential Case I ($\rho < 0$) ($\alpha = 1$, $\gamma = 0.3$, $\delta = 25$).

n	Parameter	Moment	MLE	SE(Moment)	SE(MLE)	95%CI (MM)	95%CI (MM)	PC
100	α'	4.990	4.990	0.217	0.217	(4.565, 5.415)	(4.565, 5.415)	-0.379
	β'	-19.231	-19.379	8.210	5.874	(-35.323, -3.139)	(-30.892, -7.865)	
	γ'	0.491	0.428	0.355	0.292	(-0.205, 1.186)	(-0.144, 1.000)	
	δ'	24.133	24.263	8.235	6.456	(7.991, 40.274)	(11.610, 36.917)	
	ρ	-0.377	-0.367	0.095	0.065	(-0.564, -0.191)	(-0.495, -0.240)	
1000	α'	4.999	4.999	0.069	0.069	(4.864, 5.134)	(4.865, 5.133)	-0.372
	β'	-19.633	-19.987	3.546	1.882	(-26.582, -12.683)	(-23.676, -16.299)	
	γ'	0.507	0.497	0.135	0.083	(0.242, 0.772)	(0.335, 0.660)	
	δ'	24.623	24.995	3.552	1.916	(17.661, 31.585)	(21.239, 28.752)	
	ρ	-0.371	-0.371	0.030	0.021	(-0.431, -0.312)	(-0.412, -0.330)	
10000	α'	5.000	5.000	0.022	0.022	(4.957, 5.042)	(4.957, 5.042)	-0.372
	β'	-19.987	-20.003	1.107	0.576	(-22.158, -17.816)	(-21.131, -18.875)	
	γ'	0.500	0.500	0.040	0.026	(0.421, 0.580)	(0.449, 0.551)	
	δ'	24.998	25.001	1.108	0.585	(22.826, 27.171)	(23.854, 26.148)	
	ρ	-0.372	-0.371	0.010	0.007	(-0.391, -0.353)	(-0.385, -0.358)	

Table 5: Simulations for Lomax($\eta' = 1, \gamma'$) model ($\alpha' = 5, \beta' = -20, \gamma' = 0.5, \delta' = 25$).

5 Application: Data set I

Considering the data set that is in Leiter and Hamdan [2], where the data is a 50-mile stretch of Interstate 95 in Prince William, Stafford and Spotsylvania counties in Eastern Virginia. The data is categorized as fatal accidents, injury accidents or property damage accidents, together with corresponding number of fatalities and injuries for period 1 January 1969 to 31 October 1970. Number of fatalities was considered as X_1 since the Fisher Dispersion Index is 1.052 and Number of injury accidents as X_2 (Fisher Dispersion Index is 1.14) with $n = 639$. The Pearson Correlation Coefficient for the data is 0.205 - indicating that β or β' should be set to 1 in the Case I sub-models.

We observe that the bounds specified in Equations 15 and 45 are not satisfied for this dataset. Consequently, the m.m.e's, particularly $\tilde{\gamma}$ and $\tilde{\gamma}'$, for both the full exponential model and the Lomax($\eta' = 1, \gamma'$) model, respectively, fail to exist, with the exception of $\tilde{\alpha}$ and $\tilde{\alpha}'$. Moreover, since $\frac{M_1}{S_{12}} < e$, the m.m.e's for the exponential Case I sub-model, specifically $\tilde{\gamma}$, do not exist. For the Lomax Case I sub-model, no explicit bounds governing the existence of m.m.e's (notably $\tilde{\gamma}'$) were derived; only Figure 6 provides a graphical reference. Therefore, the existence of $\tilde{\gamma}'$ in this case is determined solely by the convergence of the optimisation routine. Finally, as established earlier, the full Lomax model lacks sufficient number of moment equations to derive m.m.e's,

explaining the absence of said estimators for this model.

Additionally, recall that for the Case III sub-models (no intercept), the condition $x_1 = 0$ implies $x_2 = 0$ with probability one. Consequently, any empirical observation satisfying $x_1 = 0$ but $x_2 > 0$ is assigned zero probability under the model, leading to a likelihood of zero and thus signaling model misspecification for such observations. This result implies that both Case III sub-models, as well as the Case V sub-models (no intercept), are unsuitable for application in this setting.

Furthermore, the ρ estimate based on the m.l.e.'s of the full Lomax model is excluded, as no computational methods exist for evaluating the correlation in Equation 39 when the discreteness assumption on η' is relaxed. In such cases, one must approximate infinite sums without the simplifying closed-form representations provided by hypergeometric functions.

The m.m.e.'s, m.l.e.'s, AIC values, and the corresponding $-2\log\Lambda$ statistics (where applicable) for the various models are reported in Table 6 and Table 7 for the exponential and Lomax models, respectively. For the Case I and Case II exponential sub-models, the $-2\log\Lambda$ values did not exceed $\chi^2_{1,0.05} = 3.84$. Similarly, the Case IV exponential sub-model's $-2\log\Lambda$ value did not exceed $\chi^2_{2,0.05} = 5.99$. These results indicate insufficient evidence to reject the null hypothesis of no difference between the respective exponential sub-models and the full exponential model.

An analogous conclusion holds for the Lomax sub-models: the $-2\log\Lambda$ value for the $\text{Lomax}(\eta' = 1, \gamma')$ model is compared against $\chi^2_{1,0.05} = 3.84$; for the Case I and Case II Lomax sub-models, the comparisons are against $\chi^2_{2,0.05} = 5.99$; and for the Case IV Lomax sub-model, the comparison is against $\chi^2_{3,0.05} = 7.815$. In all cases, the corresponding $-2\log\Lambda$ values fail to exceed the critical thresholds, implying no significant difference between the full model and the sub-models.

The best-fitting models, according to the AIC criterion, are the exponential Case IV sub-model and the Lomax Case II sub-model. However, neither of these models outperforms Arnold and Manjunath [1]'s Bivariate Pseudo-Poisson Sub-Model I (A&M BPP SM-I) - a two-parameter model - with an AIC value of 1866.094. Figure 8 depicts the regression functions $E(X_2 | X_1 = x_1)$ against $X_1 = x_1$ for the aforementioned sub-models, as well as for A&M BPP SM-I.

Model	Parameter	m.m.e	m.l.e	$-2\log\Lambda$	AIC
Full exponential	α	0.058	0.058	-	1869.831
	β	-	1.240		
	γ	-	1.256		
	δ	-	0.813		
	ρ	-	0.215		
Case I ($\beta = 1$)	α	0.058	0.058	0.04	1867.872
	γ	-	2.175		
	δ	-	0.813		
	ρ	-	0.213		
Case II ($\gamma = 1$)	α	0.058	0.058	≈ 0	1867.829
	β	1.423	1.296		
	δ	0.811	0.813		
	ρ	0.219	0.215		
Case IV ($\beta = 1$ & $\gamma = 1$)	α	0.058	0.058	1.437	1867.268
	δ	0.826	0.821		
	ρ	0.156	0.156		

Table 6: m.m.e's, m.l.e's, $2\log\Lambda$'s (where applicable) and AIC's for exponential model and sub-models.

Model	Parameter	m.m.e	m.l.e	$-2 \log \Lambda$	AIC
Full Lomax	α'	0.058	0.058	-	1871.829
	β'	-	1.775		
	γ'	-	1.100		
	δ'	-	0.813		
	η'	-	1.066		
	ρ	-	-		
$\text{Lomax}(\eta' = 1, \gamma')$	α'	0.058	0.058	0.004	1869.833
	β'	-	1.533		
	γ'	-	0.726		
	δ'	-	0.813		
	ρ	-	0.215		
Case I ($\beta' = 1$)	α'	0.058	0.058	0.066	1867.895
	γ'	0.091	0.121		
	δ'	0.811	0.813		
	ρ	0.219	0.214		
Case II ($\gamma' = 1$)	α'	0.058	0.058	≈ 0	1867.829
	β'	1.803	1.769		
	δ'	0.811	0.813		
	ρ	0.219	0.215		
Case IV ($\beta' = 1 \& \gamma' = 1$)	α'	0.058	0.058	3.581	1869.409
	δ'	0.834	0.826		
	ρ	0.124	0.124		

Table 7: m.m.e's, m.l.e's, $2 \log \Lambda$'s (where applicable) and AIC's for the Lomax model and sub-models.

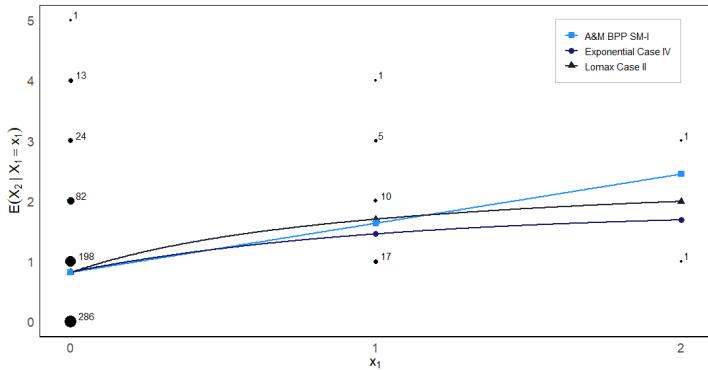


Figure 8: $E(X_2 | X_1 = x_1)$ vs. $X_1 = x_1$ for best fitted models using their m.l.e's. Size of black dots represent relative frequency of observation.

5.1 Mirrored

After assuming $X_1 \sim \text{Poisson}(\lambda_1)$ and $X_2 | X_1 \sim \text{Poisson}(\lambda_2(x_1))$, it is natural to think X_1 influences X_2 in some way. However, we entertain the possibility that the data may be better modelled by the corresponding "mirrored" model where $X_2 \sim \text{Poisson}(\lambda_1)$ and $X_1 | X_2 \sim \text{Poisson}(\lambda_2(x_2))$, just as was undergone in Arnold and Manjunath [1]. Utilizing the mirrored models results in Number of Injury Accidents being X_1 (Fisher Dispersion Index of 1.41) and Number of Fatalities being X_2 (Fisher Dispersion Index of 1.052).

We note again that, since the bounds in Equations 15 and 45 are not satisfied for this mirrored dataset, the m.m.e.'s (particularly $\tilde{\gamma}$ and $\tilde{\gamma}'$) for the mirrored full exponential model and the mirrored Lomax($\eta' = 1, \gamma'$) model do not exist, with the exception of $\tilde{\alpha}$ and $\tilde{\alpha}'$. Because the m.m.e.'s for the mirrored Case III sub-models are derived in the same manner, they are likewise omitted. Furthermore, we exclude the m.m.e.'s for the mirrored Case IV sub-models, as these yield $\tilde{\delta} < 0$ and $\tilde{\delta}' < 0$, implying said sub-models are not appropriate for this dataset.

For the mirrored dataset, we observe that whenever $X_1 = 0$, it follows that $X_2 = 0$, and never $X_2 > 0$. Consequently, it is expected that the models identify a no-intercept parameterization as appropriate, yielding estimates $\tilde{\delta} = 0$ or $\tilde{\delta}' = 0$, as consistently observed in Table 8 and Table 9.

The m.m.e.'s, m.l.e.'s, AIC values, and the corresponding $-2 \log \Lambda$ statistics (where applicable) for the various mirrored models are presented in Table 8 and Table 9. For the mirrored Case I, Case II, and Case III exponential sub-models, the $-2 \log \Lambda$ values did not exceed $\chi^2_{1,0.05} = 3.84$. Similarly, the $-2 \log \Lambda$ value for the mirrored Case V exponential sub-model did not exceed $\chi^2_{2,0.05} = 5.99$. These results again provide insufficient evidence to reject the null hypothesis of no difference between said mirrored exponential sub-models and the mirrored full exponential model.

An analogous conclusion applies to the mirrored Lomax sub-models: the $-2 \log \Lambda$ statistic for the mirrored $\text{Lomax}(\eta' = 1, \gamma')$ model is compared to $\chi^2_{1,0.05} = 3.84$; for the mirrored Case I and Case II Lomax sub-models, the comparisons are made against $\chi^2_{2,0.05} = 5.99$; and for the mirrored Case V Lomax sub-model, the comparison is made against $\chi^2_{3,0.05} = 7.815$. In all instances, the $-2 \log \Lambda$ values fail to exceed the respective critical values, indicating no significant difference between the mirrored full model and the mirrored sub-models.

We note, however, that both mirrored Case IV sub-models differ significantly from their respective mirrored full models, as indicated by their $-2 \log \Lambda$ values. This conclusion is further supported by the substantially higher AIC values observed for these mirrored sub-models relative to their corresponding mirrored full models.

Finally, we note that the mirrored exponential and mirrored Lomax Case II, Case III, and Case V sub-models achieve lower AIC values than Arnold and Manjunath [1]'s Bivariate Pseudo-Poisson Mirrored Sub-Model II ($AIC = 1847.505$) - the best-fitting model reported by Arnold and Manjunath [1] for Leiter and Hamdan [2]'s accident and fatality dataset. Figure 9 presents the regression forms $E(X_2 | X_1 = x_1)$ for both the mirrored exponential and mirrored Lomax Case V sub-models, as well as for A&M BPP MSM-II. The mirrored exponential and mirrored Lomax sub-models exhibit superior fit, which we attribute to the greater curvature of their regression forms - illustrated in Figure 9 - relative to Arnold and Manjunath [1]'s linear model.

Furthermore, the bivariate probability mass functions for the best-fitting models (both non-mirrored and mirrored) - computed using Equation 3 for the exponential model and Equation 30 for the Lomax model - are presented in Figure 10².

²Interactive content: this figure contains an embedded GIF. To view the animation, please open the PDF with **Adobe Acrobat Reader**.

Model	Parameter	m.m.e	m.l.e	$-2 \log \Lambda$	AIC
Full exponential	α	0.862	0.862	-	1848.786
	β	-	0.161		
	γ	-	0.772		
	δ	-	≈ 0		
	ρ	-	0.202		
Case I ($\beta = 1$)	α	0.862	0.862	2.082	1848.868
	γ	0.064	0.073		
	δ	0.006	≈ 0		
	ρ	0.221	0.245		
Case II ($\gamma = 1$)	α	0.862	0.862	0.082	1846.868
	β	0.159	0.143		
	δ	≈ 0	≈ 0		
	ρ	0.218	0.194		
Case III ($\delta = 0$)	α	0.862	0.862	0.131	1846.918
	β	-	0.134		
	γ	-	1.105		
	ρ	-	0.191		
Case IV ($\beta = 1 \& \gamma = 1$)	α	0.862	0.862	300.679	2145.465
	δ	-	≈ 0		
	ρ	-	0.455		
Case V ($\delta = 0 \& \gamma = 1$)	α	0.862	0.862	0.056	1844.842
	β	0.138	0.143		
	ρ	0.191	0.194		

Table 8: m.m.e's, m.l.e's, $2 \log \Lambda$'s (where applicable) and AIC's for mirrored exponential model and sub-models.

Model	Parameter	m.m.e	m.l.e	$-2 \log \Lambda$	AIC
Full Lomax	α'	0.862	0.862	-	1850.802
	β'	-	0.275		
	γ'	-	0.897		
	δ'	-	≈ 0		
	η'	-	0.510		
	ρ	-	-		
Lomax($\eta' = 1, \gamma'$)	α'	0.862	0.862	≈ 0	1848.793
	β'	-	0.217		
	γ'	-	1.490		
	δ'	-	≈ 0		
	ρ	-	0.202		
Case I ($\beta' = 1$)	α'	0.862	0.862	1.545	1848.347
	γ'	14.497	12.765		
	δ'	0.005	≈ 0		
	ρ	0.219	0.240		
Case II ($\gamma' = 1$)	α'	0.862	0.862	0.066	1846.868
	β'	0.203	0.181		
	δ'	≈ 0	≈ 0		
	ρ	0.218	0.194		
Case III ($\delta' = 0$)	α'	0.862	0.862	≈ 0	1846.793
	β'	-	0.217		
	γ'	-	1.491		
	ρ	-	0.202		
Case IV ($\beta' = 1 \& \gamma' = 1$)	α'	0.862	0.862	207.514	2052.317
	δ'	-	≈ 0		
	ρ	-	0.414		
Case V ($\delta' = 0 \& \gamma' = 1$)	α'	0.862	0.862	0.065	1844.868
	β'	0.176	0.181		
	ρ	0.190	0.194		

Table 9: m.m.e's, m.l.e's, $2 \log \Lambda$'s (where applicable) and AIC's for the mirrored Lomax model and sub-models.

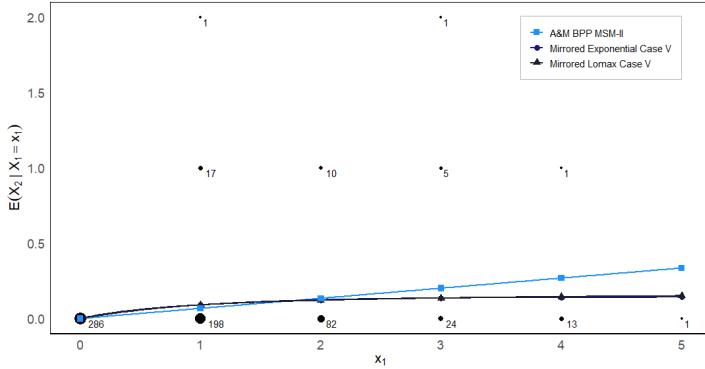


Figure 9: $E(X_2 | X_1 = x_1)$ vs. $X_1 = x_1$ for best fitted models using their m.l.e's. Size of black dots represent relative frequency of observation.

Figure 10: Mass functions of best fitted models including Accident and Fatality Data.

6 Application: Data set II

We consider the data set that is used in Islam et al. [4], where the source of the data is from the tenth wave of Health and Retirement Study (HRS). Here, X_1 represents the number of conditions ever had which has a Fisher Dispersion Index of 0.801, and X_2 represents utilization of healthcare services which has a Fisher Dispersion Index of 1.03. The Pearson Correlation Coefficient for the data is 0.063.

The m.m.e.'s, m.l.e.'s, AIC values, and the corresponding $-2 \log \Lambda$ statistics (where applicable) for the various models are reported in Table 10 and Table 11. Table 10 shows that all exponential sub-models differ significantly from the full exponential model, as evidenced by their respective $-2 \log \Lambda$ values, with worsened AIC values confirming the inferior fit of these sub-models relative to the full model. A similar conclusion follows from Table 11, where all Lomax sub-models - except for the Lomax($\eta' = 1, \gamma'$) and Lomax Case III sub-models - exhibit $-2 \log \Lambda$ values indicating significant differences from the full Lomax model.

For this dataset, we observe that whenever $X_1 = 0$, it follows that $X_2 = 0$, and never $X_2 > 0$. Consequently, the models naturally identify a no-intercept parameterization as appropriate, yielding estimates $\tilde{\delta} = 0$ or $\tilde{\delta}' = 0$.

The best-fitting models, according to the AIC criterion, are the full exponential model and the Lomax Case III sub-model, as shown in Figure 11. For comparison, Arnold and Manjunath [1]’s best-fitting model - the Bivariate Pseudo-Poisson Full Model (BPP FM) with an AIC value of 32772.08 - is also displayed. Notably, the full exponential model yields $\hat{\gamma} \approx \infty$, while the Lomax Case III sub-model yields $\hat{\gamma}' \approx 0$. This result implies that the two models are nearly identical, consistent with the previously established theoretical equivalence: where $\lim_{\hat{\gamma} \rightarrow \infty}$ in the exponential model is equivalent to $\lim_{\hat{\gamma}' \rightarrow 0}$ in the Lomax model. Theoretically it was established that such limits would produce independence models where X_2 is no longer conditioned on X_1 , since $X_1 \sim \text{Poisson}(\hat{\alpha})$ and $X_2 | X_1 \sim \text{Poisson}(\hat{\delta} + \hat{\beta})$. However, this independence only arises when $\hat{\gamma} \rightarrow \infty$ or $\hat{\gamma}' \rightarrow 0$ exactly. Empirically, we actually computed $\hat{\gamma} \approx 2 \times 10^6$ and $\hat{\gamma}' \approx 7 \times 10^{-6}$. These values indicate that, at $x_1 = 0$, the regression function is $E(X_2 | X_1 = x_1) \approx \hat{\delta}$, rather than $E(X_2 | X_1 = x_1) = \hat{\delta} + \hat{\beta}$ under the theoretical limits. Figure 11 confirms this result, showing $E(X_2 | X_1 = x_1) \approx \hat{\delta}$ (with $\hat{\delta} = 0$) for both the full exponential model and the Lomax Case III sub-model when $x_1 = 0$. For $x_1 = 1, \dots, 8$, the regression function appears constant with $E(X_2 | X_1 = x_1) \approx \hat{\delta} + \hat{\beta}$, illustrating the models’ ability to act as a linear model while correctly handling the $(x_1, x_2) = (0, 0)$ observations.

Furthermore, all models reported in Table 10 and Table 11 outperform the Bivariate COM-Poisson (BCMP) model of Arnold and Manjunath [1], which was their best-fitting model for the HRS dataset with an AIC value of 32690.18. We attribute this improvement to the exponential and Lomax models’ capacity to accurately accommodate the $(x_1, x_2) = (0, 0)$ observations.

Model	Parameter	m.m.e	m.l.e	$-2 \log \Lambda$	AIC
Full exponential	α	2.643	2.643	-	32326.411
	β	-	0.813		
	γ	-	$\approx \infty$		
	δ	-	≈ 0		
	ρ	-	0.105		
Case I ($\beta = 1$)	α	2.643	2.643	136.976	32461.387
	γ	0.032	1.238		
	δ	0.689	≈ 0		
	ρ	0.054	0.185		
Case II ($\gamma = 1$)	α	2.643	2.643	118.883	32443.294
	β	0.258	0.924		
	δ	0.560	≈ 0		
	ρ	0.057	0.199		
Case III ($\delta = 0$)	α	2.643	2.643	95.285	32419.696
	β	-	0.905		
	γ	-	1.105		
	ρ	-	0.187		
Case IV ($\beta = 1 \& \gamma = 1$)	α	2.643	2.643	146.130	32468.541
	δ	≈ 0	≈ 0		
	ρ	0.212	0.206		
Case V ($\delta = 0 \& \gamma = 1$)	α	2.643	2.643	118.883	32441.294
	β	0.947	0.924		
	ρ	0.201	0.199		

Table 10: m.m.e’s, m.l.e’s, $2 \log \Lambda$ ’s (where applicable) and AIC’s for exponential model and sub-models.

Model	Parameter	m.m.e	m.l.e	$-2 \log \Lambda$	AIC
Full Lomax	α'	2.643	2.643	-	32328.43
	β'	-	0.813		
	γ'	-	0.002		
	δ'	-	≈ 0		
	η'	-	24.324		
	ρ	-	-		
Lomax($\eta' = 1, \gamma'$)	α'	2.643	2.643	≈ 0	32326.411
	β'	-	0.813		
	γ'	-	≈ 0		
	δ'	-	≈ 0		
	ρ	-	0.105		
Case I ($\beta = 1$)	α'	2.643	2.643	53.581	32378.013
	γ'	≈ 0	0.385		
	δ'	≈ 0	≈ 0		
	ρ	0.127	0.167		
Case II ($\gamma' = 1$)	α'	2.643	2.643	146.594	32471.027
	β'	0.289	1.161		
	δ'	0.582	≈ 0		
	ρ	0.057	0.222		
Case III ($\delta' = 0$)	α'	2.643	2.643	≈ 0	32324.411
	β'	-	0.813		
	γ'	-	≈ 0		
	ρ	-	0.105		
Case IV ($\beta' = 1 \& \gamma' = 1$)	α'	2.643	2.643	174.438	32496.871
	δ'	0.121	0.088		
	ρ	0.191	0.195		
Case V ($\delta' = 0 \& \gamma' = 1$)	α'	2.643	2.643	146.594	32469.03
	β'	1.186	1.161		
	ρ	0.224	0.222		

Table 11: m.m.e's, m.l.e's, $-2 \log \Lambda$'s (where applicable) and AIC's for Lomax model and sub-models.

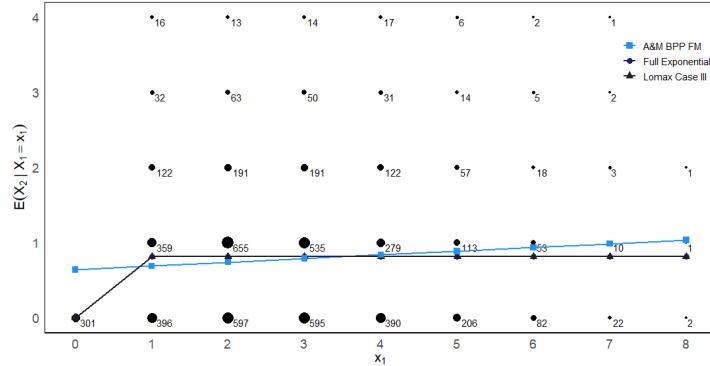


Figure 11: $E(X_2 | X_1 = x_1)$ vs. $X_1 = x_1$ for best fitted models using their m.l.e's. Size of black dots represent relative frequency of observation.

6.1 Mirrored

The m.m.e.'s, m.l.e.'s, AIC values, and the corresponding $-2 \log \Lambda$ statistics (where applicable) for the various mirrored models are reported in Table 12 and Table 13. Figure 12 presents the best-fitting models: the mirrored exponential Case I sub-model and the mirrored Lomax Case I sub-model, alongside Arnold and Manjunath [1]'s Bivariate Pseudo-Poisson Mirrored Full Model (BPP MFM) with an AIC value of 32783.080, which was the best-fitting model for the mirrored dataset in Arnold and Manjunath [1]. Figure 12 further shows that this study's best-fitting models are nearly identical to Arnold and Manjunath [1]'s BPP MFM, a linear model, indicating that the exponential and Lomax models considered in this study can approximate linear behavior when appropriate. This, in turn, suggests that a linear specification provides an adequate fit for the mirrored dataset.

Additionally, the bivariate probability mass functions for the best-fitting models - computed using Equation 3 for the exponential model and Equation 30 for the Lomax model - are presented for the HRS dataset in Figure 10.

Model	Parameter	m.m.e	m.l.e	$-2 \log \Lambda$	AIC
Full exponential	α	0.769	0.769	-	32784.617
	β	-	0.501		
	γ	-	0.287		
	δ	-	2.555		
	ρ	-	0.056		
Case I ($\beta = 1$)	α	0.769	0.769	0.144	32782.761
	γ	0.126	0.121		
	δ	2.556	2.559		
	ρ	0.058	0.056		
Case II ($\gamma = 1$)	α	0.769	0.769	1.227	32783.844
	β	0.271	0.240		
	δ	2.538	2.551		
	ρ	0.057	0.050		
Case IV ($\beta = 1 \& \gamma = 1$)	α	0.769	0.769	165.397	32946.015
	δ	2.258	2.298		
	ρ	0.205	0.203		

Table 12: m.m.e's, m.l.e's, $2 \log \Lambda$'s (where applicable) and AIC's for mirrored exponential model and sub-models.

Model	Parameter	m.m.e	m.l.e	$-2 \log \Lambda$	AIC
Full Lomax	α'	0.769	0.769	-	32786.649
	β'	-	1.517		
	γ'	-	3.746		
	δ'	-	2.556		
	η'	-	0.361		
	ρ	-	-		
Lomax($\eta' = 1, \gamma'$)	α'	0.769	0.769	≈ 0	32784.639
	β'	-	0.869		
	γ'	-	5.980		
	δ'	-	2.556		
	ρ	-	0.056		
Case I ($\beta' = 1$)	α'	0.769	0.769	≈ 0	32782.647
	γ'	7.052	7.181		
	δ'	2.555	2.556		
	ρ	0.057	0.056		
Case II ($\gamma' = 1$)	α'	0.769	0.769	1.319	32783.967
	β'	0.346	0.305		
	δ'	2.538	2.551		
	ρ	0.057	0.050		
Case IV ($\beta' = 1 \& \gamma' = 1$)	α'	0.769	0.769	86.175	32866.824
	δ'	2.340	2.364		
	ρ	0.162	0.161		

Table 13: m.m.e's, m.l.e's, $2 \log \Lambda$'s (where applicable) and AIC's for mirrored Lomax model and sub-models.

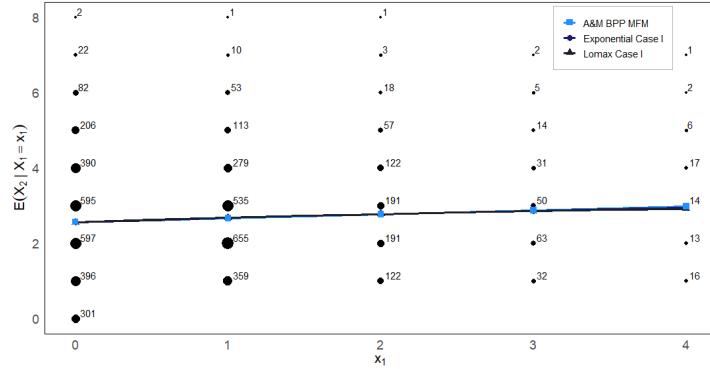


Figure 12: $E(X_2 | X_1 = x_1)$ vs. $X_1 = x_1$ for best fitted models using their m.l.e's. Size of black dots represent relative frequency of observation.

Figure 13: Mass functions of best fitted models including HRS Data.

7 Conclusion

This study set out to develop models capable of representing negative correlation between X_1 and X_2 by introducing curvature into the conditional rate $\lambda_2(x_1)$. On the HRS dataset, the proposed models modeled the boundary observation $(x_1, x_2) = (0, 0)$ more adequately than Arnold and Manjunath [1]'s linear models, while retaining the capacity to approximate linear behavior where appropriate. Additionally, the accident and fatality data indicated why added curvature is deemed necessary and may lead to improved model fit.

Caveats arose from the additional complexity of this curvature - including imposed theoretical bounds for the study's models' correlations, and data specifications for m.m.e existence. The application results further revealed that parameter estimation frequently led to near-equivalence across models - as evidenced in their respective regression plots.

Future work should assess the study's proposed models performance on negatively correlated bivariate count data and incorporate formal goodness-of-fit testing as extensively covered by Veeranna et al. [5].

Appendix

Theorem 3. $f(\tilde{\nu}) = \frac{\tilde{\alpha}^2(\tilde{\nu}-1)^2}{\tilde{\mu}(\tilde{\nu}-1)^2 - 1}$ is strictly increasing for $\tilde{\nu} \in (0, 1)$.

Proof. Note the derivative of $f(\tilde{\nu})$ with respect to $\tilde{\nu}$:

$$\frac{\partial}{\partial \tilde{\nu}} \left(\frac{\tilde{\alpha}^2(\tilde{\nu}-1)^2}{e^{\tilde{\alpha}(\tilde{\nu}-1)^2} - 1} \right) = \frac{2\tilde{\alpha}^2(\tilde{\nu}-1) \left(e^{\tilde{\alpha}(\tilde{\nu}-1)^2} (1 - \tilde{\alpha}(\tilde{\nu}-1)^2) - 1 \right)}{\left(e^{\tilde{\alpha}(\tilde{\nu}-1)^2} - 1 \right)^2}.$$

Since $0 < \tilde{\nu} < 1 \implies -2\tilde{\alpha}^2 < 2\tilde{\alpha}^2(\tilde{\nu}-1) < 0$. It remains to be shown that $f(t) = e^t(1-t) - 1 < 0$ for $t = \tilde{\alpha}(\tilde{\nu}-1)^2$ noting $0 < t < \alpha$, to ultimately show $\frac{\partial f(\tilde{\nu})}{\partial \tilde{\nu}} > 0$. Now since $\frac{d}{dt}(e^t(1-t) - 1) = -te^t$, $f(t)$ is monotonic for $t > 0$ and since $\lim_{t \rightarrow 0^+} f(t) = 0$, we have shown $f(t) < 0$ for $t > 0$. It follows that since $\frac{\partial f(\tilde{\nu})}{\partial \tilde{\nu}} > 0$ for $0 < \tilde{\nu} < 1$, $f(\tilde{\nu})$ is strictly increasing on this interval. \square

Theorem 4. $f(\tilde{\gamma}') = \frac{(\tilde{\alpha}')^2 \left[{}_1F_1(\tilde{\gamma}') - \left(\frac{\tilde{\gamma}'}{\tilde{\gamma}' + 1} \right) {}_1F_1(\tilde{\gamma}' + 1) \right]^2}{\tilde{\mu}' {}_2F_2(\tilde{\gamma}') - [{}_1F_1(\tilde{\gamma}')]^2}$ is strictly increasing for $\tilde{\gamma}' > 0$.

Proof. We aim to show $\frac{\partial f(\tilde{\gamma}')}{\partial \tilde{\gamma}'} > 0$ for $\tilde{\gamma}' > 0$. Now, through substitution of Equations 40, 41 and 41 into $f(\tilde{\gamma}')$, we have $f(\tilde{\gamma}') = f(\tilde{\nu})$ (where $f(\tilde{\nu})$ is referenced from Theorem 3). Subsequently, Theorem 3 proved $\frac{\partial f(\tilde{\nu})}{\partial \tilde{\nu}} > 0$ for $0 < \tilde{\nu} < 1$ (or equivalently, for $\tilde{\gamma} > 0$). Now since:

$$\begin{aligned} \frac{\partial f(\tilde{\gamma}')}{\partial \tilde{\gamma}'} &= \frac{\partial f(\tilde{\gamma}')}{\partial \tilde{\nu}} \cdot \frac{\partial \tilde{\nu}}{\partial \tilde{\gamma}'} \\ &= \frac{\partial f(\tilde{\nu})}{\partial \tilde{\nu}} \cdot \frac{\partial \tilde{\nu}}{\partial \tilde{\gamma}'}, \end{aligned}$$

it remains to be shown that $\frac{\partial \tilde{\nu}}{\partial \tilde{\gamma}'} > 0$. Now from Equation 40 and noting we already established $\tilde{\alpha} = \tilde{\alpha}'$, we have:

$$\begin{aligned} \tilde{\nu} &= \frac{1}{\tilde{\alpha}} \log \left({}_1F_1(\tilde{\gamma}', \tilde{\gamma}' + 1; \tilde{\alpha}') \right) \\ &= \frac{1}{\tilde{\alpha}} \log \left(\sum_{i=0}^{\infty} \frac{\tilde{\gamma}'}{\tilde{\gamma}' + i} \frac{\tilde{\alpha}'}{i!} \right). \end{aligned}$$

Now, taking the derivative of $\tilde{\nu}$ with respect to $\tilde{\gamma}'$, we have:

$$\begin{aligned} \frac{\partial \tilde{\nu}}{\partial \tilde{\gamma}'} &= \frac{1}{\tilde{\alpha}} \frac{1}{\sum_{i=0}^{\infty} \frac{\tilde{\gamma}'}{\tilde{\gamma}' + i} \frac{\tilde{\alpha}'}{i!}} \sum_{i=0}^{\infty} \frac{i}{(\tilde{\gamma}' + i)^2} \frac{\tilde{\alpha}'}{i!} \\ &> 0. \end{aligned}$$

\square

Lemma 1. $\lim_{\zeta \rightarrow 0} {}_\psi F_\psi (\zeta, \dots, \zeta, (\zeta + 1), \dots, (\zeta + 1); \xi) = 1$ and $\lim_{\zeta \rightarrow 0} {}_\psi F_\psi ((\zeta + 1), \dots, (\zeta + 1), (\zeta + 2), \dots, (\zeta + 2); \xi) = \frac{1}{\xi} (e^\xi - 1)$ for $\zeta, \xi > 0$ and $\psi \in \mathbb{N}$.

Proof. We note:

$$\begin{aligned}
\lim_{\zeta \rightarrow 0} \{ {}_\psi F_\psi (\zeta, \dots, \zeta, (\zeta + 1), \dots, (\zeta + 1); \xi) \} &= \lim_{\zeta \rightarrow 0} \left\{ \sum_{i=0}^{\infty} \left[\frac{\zeta}{\zeta + i} \right]^\psi \frac{\xi^i}{i!} \right\} \\
&= \lim_{\zeta \rightarrow 0} \left\{ 1 + \left(\frac{\zeta}{\zeta + 1} \right)^\psi \xi + \left(\frac{\zeta}{\zeta + 2} \right)^\psi \frac{\xi^2}{2} + \left(\frac{\zeta}{\zeta + 3} \right)^\psi \frac{\xi^3}{6} + \dots \right\} \\
&= 1,
\end{aligned}$$

and:

$$\begin{aligned}
\lim_{\zeta \rightarrow 0} \{ {}_\psi F_\psi ((\zeta + 1), \dots, (\zeta + 1), (\zeta + 2), \dots, (\zeta + 2); \xi) \} &= \lim_{\zeta \rightarrow 0} \left\{ \sum_{i=0}^{\infty} \left[\frac{\zeta + 1}{\zeta + 1 + i} \right]^\psi \frac{\xi^i}{i!} \right\} \\
&= \lim_{\zeta \rightarrow 0} \left\{ 1 + \left(\frac{\zeta + 1}{\zeta + 2} \right)^\psi \xi + \left(\frac{\zeta + 1}{\zeta + 3} \right)^\psi \frac{\xi^2}{2} + \left(\frac{\zeta + 1}{\zeta + 4} \right)^\psi \frac{\xi^3}{6} + \dots \right\} \\
&= \lim_{\zeta \rightarrow 0} \left\{ 1 + \left(\frac{1}{2} \right)^\psi \xi + \left(\frac{1}{3} \right)^\psi \frac{\xi^2}{2} + \left(\frac{1}{4} \right)^\psi \frac{\xi^3}{6} + \dots \right\} \\
&= \lim_{\zeta \rightarrow 0} \left\{ \sum_{i=0}^{\infty} \frac{1}{1+i} \frac{\xi^i}{i!} \right\} \\
&= \lim_{\zeta \rightarrow 0} \left\{ \frac{1}{\xi} \sum_{i=0}^{\infty} \frac{\xi^{i+1}}{(i+1)!} \right\} \\
&= \lim_{\zeta \rightarrow 0} \left\{ \frac{1}{\xi} \left(\sum_{i=0}^{\infty} \frac{\xi^i}{i!} - 1 \right) \right\} \\
&= \frac{1}{\xi} (e^\xi - 1).
\end{aligned}$$

□

$$\textbf{Theorem 5. } \lim_{\tilde{\gamma}' \rightarrow 0^+} \left\{ \frac{(\tilde{\alpha}')^2 \left[{}_1 F_1(\tilde{\gamma}') - \left(\frac{\tilde{\gamma}'}{\tilde{\gamma}' + 1} \right) {}_1 F_1(\tilde{\gamma}' + 1) \right]^2}{\tilde{\mu}' {}_2 F_2(\tilde{\gamma}') - [{}_1 F_1(\tilde{\gamma}')]^2} \right\} = \frac{(\tilde{\alpha}')^2}{\tilde{\mu}' - 1} \text{ for } \tilde{\alpha}' > 0.$$

Proof. We note:

$$\begin{aligned}
&\lim_{\tilde{\gamma}' \rightarrow 0^+} \left\{ \frac{(\tilde{\alpha}')^2 \left[{}_1 F_1(\tilde{\gamma}') - \left(\frac{\tilde{\gamma}'}{\tilde{\gamma}' + 1} \right) {}_1 F_1(\tilde{\gamma}' + 1) \right]^2}{\tilde{\mu}' {}_2 F_2(\tilde{\gamma}') - [{}_1 F_1(\tilde{\gamma}')]^2} \right\} \\
&= \frac{(\tilde{\alpha}')^2 \left[\lim_{\tilde{\gamma}' \rightarrow 0^+} \{ {}_1 F_1(\tilde{\gamma}') \} - \left(\lim_{\tilde{\gamma}' \rightarrow 0^+} \left\{ \frac{\tilde{\gamma}'}{\tilde{\gamma}' + 1} \right\} \right) \left(\lim_{\tilde{\gamma}' \rightarrow 0^+} \{ {}_1 F_1(\tilde{\gamma}' + 1) \} \right) \right]^2}{\tilde{\mu}' \lim_{\tilde{\gamma}' \rightarrow 0^+} \{ {}_2 F_2(\tilde{\gamma}') \} - [\lim_{\tilde{\gamma}' \rightarrow 0^+} \{ {}_1 F_1(\tilde{\gamma}') \}]^2},
\end{aligned}$$

and by using Lemma 1 through direct substitution, we obtain $\frac{(\tilde{\alpha}')^2}{\tilde{\mu}' - 1}$. □

References

- [1] Barry C Arnold and BG Manjunath. Statistical inference for distributions with one poisson conditional. *Journal of Applied Statistics*, 48(13-15):2306–2325, 2021.
- [2] Robert E Leiter and MA Hamdan. Some bivariate probability models applicable to traffic accidents and fatalities. *International Statistical Review/Revue Internationale de Statistique*, 41(1):87–100, 1973.

- [3] Barry C Arnold and Bangalore G Manjunath. Pseudo-poisson distributions with concomitant variables. *Mathematical and Computational Applications*, 28(1):11, 2023.
- [4] M Ataharul Islam, Rafiqul I Chowdhury, et al. *Analysis of repeated measures data*. Springer, 2017.
- [5] Banoth Veeranna, BG Manjunath, and B Shobha. Goodness of fit tests for the pseudo-poisson distribution. *arXiv preprint arXiv:2306.04168*, 2023.