مدل‌های اثر تصادفی دومتغیره آماسیده برای پاسخ‌های آمیخته سری توانی نرمال

معصومه اسمعیل‌زاده، احسان بهرامی سامانی

گروه آمار، دانشکده علوم ریاضی، دانشگاه شهید بهشتی

چکیده: این مقاله به تحلیل داده‌های آمیخته دومتغیره آماسیده می‌پردازد. برآورد پارامترهای مدل‌های مورد نظر توسط روش ماکسیمم درستنمایی انجام شده است. برای متغیر پاسخ شمارشی دومتغیره که در یک یا دو نقطه آماسیده شده، توزیع‌های جدید سری توانی آماسیده دومتغیره ارائه گردیده است. این توزیع‌های آماسیده در مدل‌بندی توأم پاسخ‌های شماشی دومتغیره مورد استفاده قرار گرفته‌اند. همچنین سودمندی مدل‌های پیشنهاد شده در چند مطالعه شبیه‌سازی بررسی و در نهایت تحلیل روی داده‌های واقعی ارائه شده است.

واژه‌های کلیدی: مدل‌های توأم، داده‌های آمیخته، مدل‌های آماسیده، اثرهای تصادفی.

۱ مقدمه

در بسیاری از مسائل کاربردی داده‌های شمارشی هستند که مقادیر آن‌ها دارای تعداد صفر زیاد است. به عنوان مثال، در پژوهش‌های آموزشی به منظور ارزیابی عملکرد دانشجویان در مطالعه بازخورد استفاده از رایانه برای بررسی سواد رایانه‌ای دانشجویان سال اول کارشناسی مد نظر است. همچنین در این پژوهش متغیرهایی همچون

°آدرس الکترونیک مسئول مقاله: احسان بهرامی سامانی، ehsan_ bahrami_ samani @yahoo.com

°کد موضوع‌بندی ریاضی (۲۰۱۰): 62J12، 62J05.

تعداد کل صفحه‌های بازدید شده و زمان کل فرایند بازدید در نظر گرفته می‌شود، به طوری که می‌توان هر دو متغیر را به عنوان متغیرهای شمارشی در نظر گرفت. مقادیر این متغیرها، ممکن است به دلیل بازدید نکردن برخی از دانشجویان دارای تعداد صفر زیادی باشند. در این مقاله سعی شده است مدل‌های مناسبی را روی این داده‌ها برازش دهیم. یکی از مهمترین مدل‌هایی که می‌توان به آن اشاره نمود مدل صفر آماسیده دو متغیره است. محققان بسیاری در مورد مدل‌های صفر آماسیده برای مدل‌بندی داده‌های شمارشی با تعداد صفر زیاد تحقیق نموده‌اند به عنوان مثال می‌توان به موارد زیر اشاره نمود: لمبرت (۱۹۹۲) اولین مدل آماری را روی داده‌های شمارشی با تعداد صفر زیاد، تحت عنوان مدل رگرسیون پواسون صفر آماسیده معرفی نمود. همچنین گوپتا و همکاران (۱۹۹۶)، ون‌دن‌براک (۱۹۹۵)، بیپ (۱۹۹۸)، زای و همکاران (۲۰۰۱)، مین و اگرستی (۲۰۰۵)، مدل‌های رگرسیون صفر آماسیده را مورد مطالعه قرار دادند. از سویی دیگر کوچرلاکوتا (۲۰۰۱) نتایج جدیدی را در مورد مدل‌های شمارشی چندمتغیره به دست آورد. از سویی دیگر پاتیل و شیرک (۲۰۱۱) مدل سری توانی دو پارامتری صفر آماسیده را مورد مطالعه قرار دادند. در این مدل‌ها، یک پارامتر برای توزیع سری توانی و یک پارامتر آمیختگی در نظر گرفته شده است. پاتیل و شیرک (۲۰۱۱) همچنین برابری پارامترهای صفر آماسیده دومتغیره سری توانی را بررسی کردند. همچنین داده‌های شمارشی دومتغیره و چندمتغیره در طیف گسترده‌ای از زمینه‌های همه‌گیرشناسی به عنوان مثال انواع مختلف یک بیماری یا بیماری در مناطق مختلف، در علوم بازاریابی و غیره کاربرد دارند. یکی از این نوع از توزیع‌ها، توزیع پواسون دو متغیره است، این توزیع توسط کوچرلاکوتا (۲۰۰۱) مورد توجه قرار گرفت. لی و همکاران (۱۹۹۹) مدل‌های مختلف پواسون صفر آماسیده چندمتغیره و ویژگی‌های توزیع‌های آن را مورد تحقیق قرار دادند. وانگ (۲۰۱۰)، یک مدل برای کاربرد داده‌های شمارشی صفر آماسیده پیشنهاد کرد و در نهایت مدل پواسون صفر آماسیده با اثرهای تصادفی به وسیله مین و اگرستی (۲۰۰۵)، اسکروندال و راب‌هسکت (۲۰۰۴) مورد تحقیق و پژوهش قرار گرفت. در نهایت رضیی و همکاران (۲۰۱۷) به بررسی و تحلیل پاسخ‌های دومتغیره آمیخته با پاسخ‌های شمارشی و ترتیبی طولی با امکان گم‌شدگی غیر قابل چشم‌پوشی پرداختند، به طوری‌که پاسخ‌های شمارشی از توزیع‌های سری توانی پیروی می‌کردند. همچنین تبریزی و همکاران (۲۰۱۸) نیز تحلیل پاسخ‌های دومتغیره شمارشی و پاسخ‌های پیوسته محدود شده در بازه (۰, ۱) انجام داده و نتایج جدیدی را به دست آوردند. دهقانی و همکاران (۱۳۹۸) به بررسی مدل‌های پواسون دو متغیره با صفر آماسیده با و بدون اثرهای تصادفی پرداختند به طوری که اثرهای تصادفی در این مدل‌ها دارای توزیع چوله نرمال هستند.

در این مقاله به معرفی مدل‌های توأم جدیدی برای تحلیل پاسخ‌های دو متغیره آماسیده پرداخته می‌شود، به طوری‌که در دو متغیر پیوسته و شمارشی، آماسیدگی و تورم در یک یا چند نقطه وجود دارد. در پژوهش‌های گذشته، مدل‌هایی که در این مقاله به آن‌ها اشاره می‌شود، مورد توجه محققان قرار نگرفته است. این مدل‌های جدید به شرح زیر است: مدل‌هایی با پاسخ‌های دومتغیره آمیخته و آماسیده در نقطه صفر برای توزیع‌های سری توانی–سری توانی و گاما–سری توانی در حالت مقطعی با استفاده از اثرهای تصادفی و تعمیم این مدل‌ها برای حالتی که آماسیدگی در دو نقطه k و ℓ وجود داشته است. لازم به ذکر است برای هر یک از پاسخ‌های پیوسته می‌توان توزیع گاما و با گسسته کردن این پاسخ توزیع سری توانی در نظر گرفت . برای پاسخ شمارشی مورد نظر نیز توزیع سری توانی مورد استفاده قرار می‌گیرد. از مهمترین این مدل‌ها می‌توان به مدل پواسون–پواسون صفر آماسیده، پواسون– گاما صفر آماسیده و درنهایت مدل دوجمله‌ای منفی–گاما صفر آماسیده اشاره نمود. در بخش ۲، مدل‌های جدید معرفی می‌شود. این مدل‌ها با عناوین مدل‌های دومتغیره مقطعی برای پاسخ‌های آمیخته سری توانی–سری توانی آماسیده در دو نقطه دلخواه k و ℓ حالت خاصی از آن مدل‌های سری توانی–سری توانی صفر آماسیده و در نهایت مدل‌هایی برای توزیع‌های سری توانی–گاما مورد بررسی قرار گرفته اند. در بخش ۳، برای تشریح سودمندی مدل‌های مطرح شده مطالعه‌های شبیه‌سازی شده صورت گرفته و در نهایت در بخش ۴ کاربرد مدل‌های مطرح شده روی داده‌های واقعی بیان شده است.

۲ مدل‌ها و تابع‌های درستنمایی

در بسیاری از مسائل کاربردی در حوزه‌های مختلفی همچون علوم پزشکی، علوم اجتماعی، علوم ورزشی و غیره پاسخ‌های آمیخته، از نوع شمارشی هستند و هدف برازش و بررسی توأم این پاسخ‌ها است. برای مثال، یک مطالعه بازخورد بر روی دانشجویان به صورت مجازی انجام شده است، پاسخ‌هایی همچون تعداد فرم‌های نظرسنجی بازدید شده در سایت توسط دانشجویان و مدت زمان لازم برای پر کردن فرم‌ها به ترتیب به عنوان پاسخ شمارشی و پیوسته، برای این افراد در نظر گرفته شده است. واضح است که بین این دو پاسخ آمیخته، همبستگی وجود دارد، زیرا هر چه تعداد فرم‌های بازدید شده بیشتر باشد زمان لازم برای پر کردن این فرم‌ها بیشتر خواهد بود.

۱.۲ تحلیل دومتغیره با پاسخ‌های آمیخته سری توانی-سری توانی آماسیده (k,ℓ)

فرض کنید Y_i^f پاسخ‌های شمارشی دارای توزیع سری توانی آماسیده (k,ℓ) است، که با نماد (ZIPS) نشان داده شده است و به صورت

$$Y_i^f \sim \begin{cases} k & \phi_۱ \quad \text{با احتمال} \\ \ell & \phi_۲\,(۱-\phi_۱) \quad \text{با احتمال} \\ PS(\vartheta_i^{(f)}) & (۱-\phi_۱)\,(۱-\phi_۲) \quad \text{با احتمال} \end{cases}$$

تعریف می‌گردد، به طوری که در آن، $PS(\vartheta_i)$ تابع جرم احتمال توزیع سری توانی است که به صورت

$$P(Y_i=j)=\frac{a(j)\vartheta_i^j}{g(\vartheta_i)},$$

است، که در آن، $a(\cdot)$ و $g(\cdot)$ به ترتیب یک تابع نامنفی و $g(\vartheta_i)=\sum_j a(j)\vartheta_i^j$. از مهمترین توزیع‌های سری‌توانی می‌توان به توزیع پواسون و دوجمله‌ای منفی اشاره نمود.

فرض کنید T_i^f پاسخ‌های پیوسته دارای توزیع سری توانی آماسیده (k,ℓ) است، که به صورت

$$P(T_i^f=k|\vartheta_i=\vartheta_i^{(t)})=\phi_۱+(۱-\phi_۱)(۱-\phi_۲)\tfrac{a(k)\vartheta_i^k}{g(\vartheta_i)}/C_i(K),$$

$$P(T_i^f=\ell|\vartheta_i=\vartheta_i^{(t)})=\phi_۲(۱-\phi_۱)+(۱-\phi_۱)(۱-\phi_۲)\tfrac{a(\ell)\vartheta_i^\ell}{g(\vartheta_i)}/C_i(K),$$

$$P(T_i^f=j|\vartheta_i=\vartheta_i^{(t)})=(۱-\phi_۱)(۱-\phi_۲)\tfrac{a(j)\vartheta_i^j}{g(\vartheta_i)}/C_i(K).$$

است، که در آن $C_i(K)$ به صورت

$$C_i(K)=P(Y_i^f\le K|\vartheta_=\vartheta_i^f)=\sum_{j=۰}^{K}\tfrac{a(j)\vartheta_i^j}{g(\vartheta_i)}$$

تعریف می‌شود. مدل دو متغیره با پاسخ‌های آماسیده (k,ℓ) به صورت

$$Y_i^f|\vartheta_i^{(f)}\sim ZIPS(\vartheta_i^{(f)},\phi_۱,\phi_۲),$$

$$T_i^f|\vartheta_i^{(t)}\sim ZIPS(\vartheta_i^{(t)},\phi_۱,\phi_۲),$$

$$q_۱(\vartheta_i^{(f)})=X'_{۱i}\beta_۱+W'_{۱i}\eta_۱,$$

$$q_۲(\phi_۱)=X'_{۲i}\beta_۲,$$

$$q_۳(\phi_۲)=X'_{۳i}\beta_۳,$$

$$q_۴(\vartheta_i^{(t)}) = X'_{۴i}\beta_۴ + W'_{۲i}\eta_۲.$$

خواهد بود، که در آن $q_۱(\cdot)$، $q_۲(\cdot)$ و $q_۳(\cdot)$ توابع ربط مناسب برای مدل‌های خطی تعمیم‌یافته مربوط به پارامترهای $\vartheta_i^{(f)}$، $\phi_۱$، $\phi_۲$ و $\vartheta_i^{(t)}$ هستند. معمولاً برای این مدل، $q_۱$ و $q_۴$ تابع پیوند لگاریتم و $q_۲$ و $q_۳$ تابع پیوند لوجیت در نظر گرفته می‌شود. همچنین $X_{۱i}$، $X_{۲i}$، $X_{۳i}$ و $X_{۴i}$ سطر iام ماتریس‌های طرح $X_۱$، $X_۲$، $X_۳$ و $X_۴$ شامل متغیرهای تبیینی هستند و بردارهای $W_{۱i}$ و $W_{۲i}$ سطر iام ماتریس‌های طرح $W_۱$ و $W_۲$ هستند به طوری که زیرماتریس‌هایی از ماتریس‌های طرح $X_۱$ ، $X_۲$ در نظر گرفته می‌شوند. همچنین $\eta_۱$ و $\eta_۲$ بردارهای اثر تصادفی هستند، به طوری که $(\eta_۱, \eta_۲)$ دارای توزیع نرمال دومتغیره با بردار میانگین صفر و ماتریس کوواریانس Σ_η است.

تابع درستنمایی مدل، به منظور برآورد پارامترهای مدل به روش ماکسیمم درستنمایی به صورت

$$\begin{aligned} L &= \prod_{i=۱}^{n} \int f(y_i^f, t_i^f|\eta_۱, \eta_۲) f(\eta_۱, \eta_۲) d\eta_۱ d\eta_۲ \\ &= \prod_{i=۱}^{n} \int f(y_i^f|\eta_۱) f(t_i^f|\eta_۲) f(\eta_۱) d\eta_۱ d\eta_۲, \end{aligned}$$

به دست می‌آید، با در نظر گرفتن تابع پیوند لوجیت برای مدل‌های مربوط به پارامتر آماسیده، خواهیم داشت:

$$\begin{aligned} f(y_i^f|\eta_۱, \eta_۲, \eta_۳) &= \big(۱ + \exp(x'_{۲i}\beta_۲)\big)^{-۱} \\ &\times \Big(\exp(x'_{۲i}\beta_۲) + \big(۱ + \exp(x'_{۳i}\beta_۳)\big)\frac{a(k) q_۱^{-۱}(x'_{۱i}\beta_۱ + w'_{۱i}\eta_۱)^k}{g\big(q_۱^{-۱}(x'_{۱i}\beta_۱ + w'_{۱i}\eta_۱)\big)}\Big)^{o_k} \\ &\times \Big(\exp(x'_{۳i}\beta_۳)\big(۱ + \exp(x'_{۳i}\beta_۳)\big)^{-۱} \\ &+ \big(۱ + \exp(x'_{۳i}\beta_۳)\big)^{-۱}\frac{a(\ell) q_۱^{-۱}(x'_{۱i}\beta_۱ + w'_{۱i}\eta_۱)^\ell}{g\big(q_۱^{-۱}(x'_{۱i}\beta_۱ + w'_{۱i}\eta_۱)\big)}\Big)^{o_\ell} \\ &\times \Big(\big(۱ + \exp(x'_{۳i}\beta_۳)\big)^{-۱}\frac{a(j) q_۱^{-۱}(x'_{۱i}\beta_۱ + w'_{۱i}\eta_۱)^j}{g\big(q_۱^{-۱}(x'_{۱i}\beta_۱ + w'_{۱i}\eta_۱)\big)}\Big)^{(۱-o_k)(۱-o_\ell)} \end{aligned}$$

$$\begin{aligned} f(t_i^f|\eta_۱, \eta_۲, \eta_۳) &= \big(۱ + \exp(x'_{۲i}\beta_۲)\big)^{-۱} \\ &\times \Big(\exp(x'_{۲i}\beta_۲) + \big(۱ + \exp(x'_{۳i}\beta_۳)\big)\frac{a(k) q_۴^{-۱}(x'_{۴i}\beta_۴ + w'_{۲i}\eta_۲)^k}{g\big(q_۴^{-۱}(x'_{۴i}\beta_۴ + w'_{۲i}\eta_۲)\big)}\Big)^{o_k} \\ &\times \Big(\exp(x'_{۳i}\beta_۳)\big(۱ + \exp(x'_{۳i}\beta_۳)\big)^{-۱} \\ &+ \big(۱ + \exp(x'_{۳i}\beta_۳)\big)^{-۱}\frac{a(\ell) q_۴^{-۱}(x'_{۴i}\beta_۴ + w'_{۲i}\eta_۲)^\ell}{g\big(q_۴^{-۱}(x'_{۴i}\beta_۴ + w'_{۲i}\eta_۲)\big)}\Big)^{o_\ell} \\ &\times \Big(\big(۱ + \exp(x'_{۳i}\beta_۳)\big)^{-۱}\frac{a(j) q_۴^{-۱}(x'_{۴i}\beta_۴ + w'_{۲i}\eta_۲ j}{g\big(q_۴^{-۱}(x'_{۴i}\beta_۴ + w'_{۲i}\eta_۲)\big)}\Big)^{(۱-o_k)(۱-o_\ell)} \end{aligned}$$

یکی از حالت های خاص مدل بیان شده، مدل سری توانی-سری توانی صفر آماسیده است که در ادامه این مدل معرفی می‌شود. پاسخ Y_i^f $(i = 1, ..., n, k = 0, ..., K)$، پاسخی شمارشی با تعداد صفر زیاد از توزیع سری توانی صفر آماسیده $(ZIPS)$ است. بنابراین توزیع متغیر تصادفی Y_i^f، آمیخته‌ای از توزیع سری توانی و یک توزیع تباهیده در نقطه صفر در نظر گرفته می‌شود. مدل سری توانی صفر آماسیده مربوط به مؤلفه iام نمونه به صورت

$$Y_i^f \sim \begin{cases} 0 & \text{با احتمال} \quad 1 - \phi_i \\ PS(\vartheta_i^{(f)}) & \text{با احتمال} \quad \phi_i \end{cases}$$

معرفی می‌شود، که در آن ϕ_i پارامتر آمیختگی مربوط به مؤلفه iام است. از سویی دیگر احتمال این‌که متغیر پاسخ iام به ترتیب یک مقدار صفر و یک مقدار غیر صفر را اختیار کند، به صورت

$$P(Y_i^f = 0|\vartheta_i = \vartheta_i^{(f)}, Y_i^f \leq K) = (1 - \phi_i) + \phi_i \tfrac{a(0)}{g(\vartheta_i)} / C_i(K),$$

$$P(Y_i^f = j|\vartheta_i = \vartheta_i^{(f)}, Y_i^f \leq K) = \phi_i \tfrac{a(j)\vartheta_i^j}{g(\vartheta_i)} / C_i(K).$$

محاسبه می‌شود. که در آن $C_i(K)$ به صورت $C_i(K) = P(Y_i^f \leq K|\vartheta_= \vartheta_i^f) = \sum_{j=0}^{K} \frac{a(j)\vartheta_i^j}{g(\vartheta_i)}$ تعریف می‌گردد، همچنین پاسخ T_i^f پاسخی از توزیع سری توانی صفر آماسیده است که نشان دهنده مدت زمان بازدید از فرم‌های نظر سنجی خواهد بود. همچنین T_i^f، آمیخته‌ای از یک توزیع تباهیده در نقطه صفر با پارامتر آمیختگی ϕ_i و یک توزیع سری توانی برای مشاهدات غیر صفر در نظر گرفته می‌شود. احتمال مشاهده مقدار صفر و غیر صفر برای پاسخ زمان به صورت

$$P(T_i^f = 0|\vartheta_i = \vartheta_i^{(t)}) = (1 - \phi_i) + \phi_i \tfrac{a(0)}{g(\vartheta_i)},$$
$$P(T_i^f = j|\vartheta_i = \vartheta_i^{(t)}) = \phi_i \tfrac{a(j)\vartheta_i^j}{g(\vartheta_i)},$$

به دست می‌آید. از سویی دیگر می‌توان یک مدل جدید دومتغیره برای دو پاسخ که دارای توزیع سری توانی صفر آماسیده است به صورت زیر معرفی نمود، این مدل به صورت

$$Y_i^f|\vartheta_i^{(f)}, \phi_i \sim ZIPS(\vartheta_i^{(f)}, \phi_i),$$
$$T_i^f|\vartheta_i^{(t)}, \phi_i \sim ZIPS(\vartheta_i^{(t)}, \phi_i),$$
$$q_1(\vartheta_i^{(f)}) = X'_{1i}\beta_1 + W'_{1i}\eta_1,$$

$$q_2(\phi_i) = X'_{2i}\beta_2,$$
$$q_3(\vartheta_i^{(t)}) = X'_{3i}\beta_3 + W'_{2i}\eta_2,$$

بیان می‌گردد، که در آن $q_1(\cdot)$ ، $q_2(\cdot)$ و $q_3(\cdot)$ توابع پیوند مناسب برای مدل‌های خطی تعمیم‌یافته مربوط به پارامترهای $\vartheta_i^{(f)}$ ، ϕ_i و $\vartheta_i^{(t)}$ هستند. معمولاً برای این مدل، q_1 و q_3 تابع پیوند لگاریتم و q_2 تابع پیوند لوجیت در نظر گرفته می‌شود. همچنین X_{1i} ، X_{2i} و X_{3i} سطر iام ماتریس‌های طرح X_1 ، X_2 و X_3 شامل متغیرهای تبیینی هستند. از سویی دیگر بردارهای W_{1i} و W_{2i} سطر iام ماتریس‌های طرح W_1 و W_2 که زیرماتریس‌هایی از ماتریس‌های طرح X_1 ، X_2 ، X_3 هستند و همچنین η_1 و η_2 بردارهای اثر تصادفی در نظر گرفته می‌شوند که (η_1, η_2) دارای توزیع نرمال دومتغیره با بردارهای میانگین صفر و ماتریس کوواریانس Σ_η است. به منظور برآورد پارامترهای مدل از روش ماکسیمم درستنمایی استفاده می‌شود. برای این منظور تابع درستنمایی مربوط به مدل به صورت

$$\begin{aligned} L &= \prod_{i=1}^{n} \int_{A_i} f(y_i^f, t_i^f | \eta_1, \eta_2).f(\eta_1, \eta_2) d\eta_1 d\eta_2 \\ &= \prod_{i=1}^{n} \int_{A_i} f(y_i^f | \eta_1).f(t_i^f | \eta_2) f(\eta_1, \eta_2) d\eta_1 d\eta_2, \end{aligned}$$

به دست می‌آید، که در آن

$$A_i = \{(\eta_1, \eta_2); \eta_1 \in R^{d_1}, \eta_2 \in R^{d_2}\}.$$

با در نظر گرفتن توابع پیوند مورد نظر، توابع جرم احتمالی که در تابع درستنمایی مدل استفاده شده است به صورت

$$\begin{aligned} f(y_i^f | \eta_1, \eta_2) &= \big(1 + \exp(x'_{2i}\beta_2)\big)^{-1} \\ &\times \Big(1 + \exp(x'_{2i}\beta_2) \frac{a(0)}{g(q_1^{-1}(x'_{1i}\beta_1 + w'_{1i}\eta_1))}\Big)^{1-o_i} \\ &\times \Big(\exp(x'_{2i}\beta_2) \frac{a(j) q_1^{-1}(x'_{1i}\beta_1 + w'_{1i}\eta_1)^j}{g(q_1^{-1}(x'_{1i}\beta_1 + w'_{1i}\eta_1))}\Big)^{o_i} \end{aligned}$$

و

$$f(t_i^f | \eta_1, \eta_2) = \big(1 + \exp(x'_{2i}\beta_2)\big)^{-1}$$

$$\times \quad \big(1+\exp(x'_{2i}\beta_2)\frac{a(0)}{g(q_3^{-1}(x'_{3i}\beta_3+w'_{2i}\eta_2))}\big)^{1-o_i},$$
$$\times \quad \big(\exp(x'_{2i}\beta_2)\frac{a(j)q_3^{-1}(x'_{3i}\beta_3+w'_{2i}\eta_2)^j}{g(q_3^{-1}(x'_{3i}\beta_3+w'_{2i}\eta_2))}\big)^{o_i},$$

محاسبه می‌گردند. از مهمترین این مدل‌ها می‌توان به مدل‌های پواسون–پواسون صفر آماسیده اشاره نمود.

۲.۲ تحلیل دو متغیره با پاسخ‌های آمیخته سری توانی-گاما آماسیده (k,ℓ)

در برخی از مطالعه‌ها، پاسخ‌های آمیخته از نوع سری توانی و پیوسته هستند، در این پاسخ‌ها ممکن است متغیرهای مورد نظر دارای ویژگی‌های آماسیده شدن در دو نقطه k و ℓ باشند. در ادامه به معرفی مدل‌هایی برای تحلیل این پاسخ‌ها پرداخته می‌شود. فرض کنید Y_i^f پاسخ شمارشی از توزیع سری توانی آماسیده (k,ℓ) است، که به صورت

$$Y_i^f \sim \begin{cases} k & \phi_1 \quad \text{با احتمال} \\ \ell & \phi_2\,(1-\phi_1) \quad \text{با احتمال} \\ PS(\vartheta_i^{(f)}) & (1-\phi_1)\,(1-\phi_2) \quad \text{با احتمال} \end{cases}$$

تعریف می‌شود. با فرض این‌که مشاهدات زمان به صورت پیوسته ثبت شوند و دارای توزیع گامای آماسیده (k,ℓ) است، خواهیم داشت:

$$P(T_i^f=k|\vartheta_i=\vartheta_i^{(t)})=\phi_1+\frac{(1-\phi_1)(1-\phi_2)}{(\frac{2\vartheta_i}{\nu})^{\frac{\nu}{2}}\Gamma(\frac{\nu}{2})}k^{\frac{\nu-2}{2}}\exp(\frac{-k\nu}{2\vartheta_i}),$$
$$P(T_i^f=\ell|\vartheta_i=\vartheta_i^{(t)})=\phi_2(1-\phi_1)+\frac{(1-\phi_1)(1-\phi_2)}{(\frac{2\vartheta_i}{\nu})^{\frac{\nu}{2}}\Gamma(\frac{\nu}{2})}\ell^{\frac{\nu-2}{2}}\exp(\frac{-\ell\nu}{2\vartheta_i}),$$
$$P(T_i^f=j|\vartheta_i=\vartheta_i^{(t)})=\frac{(1-\phi_1)(1-\phi_2)}{(\frac{2\vartheta_i}{\nu})^{\frac{\nu}{2}}\Gamma(\frac{\nu}{2})}j^{\frac{\nu-2}{2}}\exp(\frac{-j\nu}{2\vartheta_i}).$$

مدل دومتغیره با پاسخ‌های آماسیده (k,ℓ) به صورت

$$Y_i^f|\vartheta_i^{(f)}\sim ZIPS(\vartheta_i^{(f)},\phi_1,\phi_2),$$
$$T_i^f|\vartheta_i^{(t)}\sim ZIG(\vartheta_i^{(t)},\phi_1,\phi_2),$$
$$q_1(\vartheta_i^{(f)})=X'_{1i}\beta_1+W'_{1i}\eta_1,$$
$$q_2(\phi_1)=X'_{2i}\beta_2,$$
$$q_3(\phi_2)=X'_{3i}\beta_3,$$
$$q_4(\vartheta_i^{(t)})=X'_{4i}\beta_4+W'_{2i}\eta_2$$

خواهد بود، که در آن $q_1(\cdot)$، $q_2(\cdot)$ و $q_3(\cdot)$ توابع پیوند مناسب برای مدل‌های خطی تعمیم‌یافته مربوط به پارامترهای $\vartheta_i^{(f)}$ ، ϕ_1 ، ϕ_2 و $\vartheta_i^{(t)}$ هستند. معمولاً برای این مدل، q_1 و q_4 تابع پیوند لگاریتم و q_2 و q_3 تابع پیوند لوجیت در نظر گرفته می‌شود. همچنین X_{1i} ، X_{2i} ، X_{3i} و X_{4i} سطر iام ماتریس‌های طرح X_1 ، X_2 ، X_3 و X_4 شامل متغیرهای تبیینی هستند و بردارهای W_{1i} و W_{2i} سطر iام ماتریس‌های طرح W_1 و W_2 که زیرماتریس‌هایی از ماتریس‌های طرح X_1 ، X_2 در نظر گرفته می‌شوند. همچنین η_1 و η_2 بردارهای اثر تصادفی هستند، به طوری که (η_1, η_2) دارای توزیع نرمال دومتغیره با بردار میانگین صفر و ماتریس کوواریانس Σ_η است. تابع درستنمایی مدل، به منظور برآورد پارامترهای مدل به روش ماکسیمم درستنمایی به صورت

$$\begin{aligned} L &= \prod_{i=1}^{n} \int f(y_i^f, t_i^f | \eta_1, \eta_2) . f(\eta_1, \eta_2) d\eta_1 d\eta_2 \\ &= \prod_{i=1}^{n} \int f(y_i^f | \eta_1) f(t_i^f | \eta_2) . f(\eta_1) d\eta_1 d\eta_2, \end{aligned}$$

به دست می‌آید. با در نظر گرفتن تابع پیوند لوجیت برای مدل‌های مربوط به پارامتر آماسیده، خواهیم داشت:

$$\begin{aligned} f(y_i^f | \eta_1, \eta_2, \eta_3) = {} & \big(1 + \exp(x'_{2i}\beta_2)\big)^{-1} \\ & \times \Big(\exp(x'_{2i}\beta_2) + \big(1 + \exp(x'_{3i}\beta_3)\big) \frac{a(k) q_1^{-1}(x'_{1i}\beta_1 + w'_{1i}\eta_1)^k}{g\big(q_1^{-1}(x'_{1i}\beta_1 + w'_{1i}\eta_1)\big)} \Big)^{o_k} \\ & \times \Big(\exp(x'_{3i}\beta_3) \big(1 + \exp(x'_{3i}\beta_3)\big)^{-1} \\ & + \big(1 + \exp(x'_{3i}\beta_3)\big)^{-1} \frac{a(\ell) q_1^{-1}(x'_{1i}\beta_1 + w'_{1i}\eta_1)^\ell}{g\big(q_1^{-1}(x'_{1i}\beta_1 + w'_{1i}\eta_1)\big)} \Big)^{o_\ell} \\ & \times \Big(\big(1 + \exp(x'_{3i}\beta_3)\big)^{-1} \frac{a(j) q_1^{-1}(x'_{1i}\beta_1 + w'_{1i}\eta_1)^j}{g\big(q_1^{-1}(x'_{1i}\beta_1 + w'_{1i}\eta_1)\big)} \Big)^{(1-o_k)(1-o_\ell)} \end{aligned}$$

و

$$\begin{aligned} f(t_i^f | \eta_2, \eta_3, \eta_4) &= \big(1 + \exp(x'_{2i}\beta_2)\big)^{-1} \Big[\exp(x'_{2i}\beta_2) \\ &+ \frac{\big(1 + \exp(x'_{3i}\beta_3)\big)}{\big(\frac{2 q_4^{-1}(x'_{4i}\beta_4 + w'_{2i}\eta_2)}{\nu}\big)^{\frac{\nu}{2}} \Gamma(\frac{\nu}{2})} k^{\frac{\nu-2}{2}} \exp\Big(\frac{-k\nu}{2 q_4^{-1}(x'_{4i}\beta_4 + w'_{2i}\eta_2)}\Big)\Big]^{o_k} \\ &\times \Big[\exp(x'_{3i}\beta_3) \big(1 + \exp(x'_{3i}\beta_3)\big)^{-1} \\ &+ \frac{\big(1 + \exp(x'_{3i}\beta_3)\big)}{\big(\frac{2 q_4^{-1}(x'_{4i}\beta_4 + w'_{2i}\eta_2)}{\nu}\big)^{\frac{\nu}{2}} \Gamma(\frac{\nu}{2})} \ell^{\frac{\nu-2}{2}} \exp\Big(\frac{-\nu}{2 q_4^{-1}(x'_{4i}\beta_4 + w'_{2i}\eta_2)}\Big)\Big]^{o_\ell} \\ &\times \Big[\frac{\big(1 + \exp(x'_{3i}\beta_3)\big)}{\big(\frac{2 q_4^{-1}(x'_{4i}\beta_4 + w'_{2i}\eta_2)}{\nu}\big)^{\frac{\nu}{2}} \Gamma(\frac{\nu}{2})} j^{\frac{\nu-2}{2}} \end{aligned}$$

$$\times \quad \exp(\frac{-j\nu}{۲q_۴^{-۱}(x'_{۴i}\beta_۴ + w'_{۲i}\eta_۲)})\Big]^{(۱-o_k)(۱-o_\ell)}.$$

یکی از حالت‌های خاص مدل بیان شده در این زیر بخش، مدل سری توانی–گاما صفر آماسیده است که همانند زیر بخش ۱.۲، مدل و تابع درستنمایی آن قابل حصول است. از مهمترین این مدل‌ها می‌توان به مدل پواسون–گاما و مدل دوجمله‌ای منفی–گاما صفر آماسیده اشاره نمود.

۳ مطالعه شبیه‌سازی

در این بخش برای تشریح سودمندی مدل‌های بیان شده، مطالعه شبیه‌سازی انجام می‌گردد. مطالعه شبیه‌سازی برای مدل‌های پواسون صفر آماسیده، مدل گاما صفر آماسیده، مدل پواسون–پواسون صفر آماسیده، مدل پواسون–گاما صفر آماسیده و مدل دو جمله‌ای منفی–گاما صفر آماسیده ارائه می‌شوند. در این مطالعه شبیه‌سازی ابتدا نحوه تولید داده‌ها بیان می‌شود و سپس مدل‌های مورد نظر روی داده‌های شبیه‌سازی شده برازش می‌شوند.

۱.۳ مدل پواسون صفر آماسیده (ZIP)

فرض می‌شود متغیر سری توانی Y_i^f دارای توزیع پواسون با پارامتر $\lambda_i^{(f)}$ است. همچنین متغیر اثر تصادفی η_i در مدل خطی برای متغیر پاسخ گسسته Y_i^f دارای توزیع نرمال با میانگین صفر و واریانس $\sigma_{\eta_i}^۲ = ۱$ و X تنها متغیر تبیینی بر پاسخ شمارشی Y_i^f در نظر گرفته شود. در همه این مدل‌ها متغیر تبیینی X_i از توزیع نرمال با میانگین صفر و واریانس یک تولید می‌گردند. مقادیر واقعی پارامترها به صورت

$$(\beta_۰, \beta_۱, \xi_۰, \xi_۱, \sigma_{\eta_i}^۲)' = (۰, ۱, ۰, ۱, ۱)',$$

در مدل‌های مورد نظر

$$Y_i^f|\lambda_i^{(f)} \sim ZIP(\lambda_i^{(f)}, \phi_i),$$
$$\log(\lambda_i^{(f)}) = \beta_۰ + \beta_۱ X_i + \sigma_\eta \eta_i,$$
$$\text{logit}(\phi_i) = \xi_۰ + \xi_۱ X_i,$$

قرار می‌گیرند. بنابراین متغیر تصادفی Y_i^f از توزیع پواسون صفر آماسیده به این شرح تولید می‌گردد: ابتدا متغیر تصادفی u_i از توزیع یکنواخت روی بازه $(۰, ۱)$ تولید شده و سپس با در نظر گرفتن $\phi_i = ۰/۸$

(احتمال آمیختگی) و مقایسه مقادیر u_i با ۰/۸ مقادیر Y_i^f به صورت زیر تولید می‌شوند که اگر $u_i \leq ۰/۸$ ، آن‌گاه $y_i^f = ۰$ ، اگر $u_i > ۰/۸$ ، آن‌گاه متغیر تصادفی Y_i^f از توزیع پواسون با پارامتر $\lambda_i^{(f)}$ تولید می‌شود. در نهایت بعد از تولید نمونه تصادفی بر اساس مدل مطرح شده، تابع درستنمایی برای این مدل به صورت

$$L = \Pi_{i=۱}^{n} \int_{\eta_i} f(y_i^f|\eta_i) f(\eta_i) d\eta_i$$

به دست می‌آید. از آن‌جایی که حل انتگرال فوق ممکن است با دشواری‌هایی روبه‌رو گردد، بنابراین با استفاده از روش تقریب انتگرال به وسیله روش $MCMC$ انتگرال فوق تقریب زده می‌شود. برای این منظور ابتدا یک نمونه تصادفی به حجم m از توزیع $\eta_i \sim N(۰, ۱)$ تولید می‌شود، سپس انتگرال مورد نظر به صورت زیر تقریب زده می‌شود

$$\int_{\eta_i} f(y_i^f|\eta_i) f(\eta_i) d\eta_i \simeq \frac{۱}{m} \Sigma_{k=۱}^{m} f(y_i^f|\eta_i^{(k)})$$

حال با روش نیوتون-رافسون منفی لگاریتم تقریب تابع درستنمایی مینیمم می‌شود و برآورد پارامترهای مدل به دست می‌آید، برای این منظور از دستور $nlminb$ در نرم افزار R استفاده می‌گردد. همچنین ماتریس اطلاع فیشر مربوط به برآورد پارامترهای مدل را به دست آورده، سپس عکس این ماتریس محاسبه شده، درایه‌های روی قطر اصلی این ماتریس، واریانس‌های مربوط به برآورد پارامترهای مدل هستند. در نهایت خطای استاندارد برآورد پارامترهای مدل از جذر مثبت واریانس‌های موجود در عکس ماتریس اطلاع فیشر به دست می‌آید. برای به دست آوردن ماتریس اطلاع فیشر از دستور $fdHess$ در نرم افزار R استفاده می‌شود. برای انجام کار شبیه‌سازی نمونه‌هایی به حجم ۵۰، ۱۰۰ و ۱۰۰۰تایی با استفاده از نرم‌افزار R از توزیع Y تولید شده است. برآورد نقطه‌ای، خطای استانداد هر یک از پارامترهای مدل برای حجم نمونه‌های ۵۰، ۱۰۰ و ۱۰۰۰ محاسبه و در جدول ۱ ارائه شده است. فرایند تولید نمونه و برآورد پارامترهای مدل در حجم‌های مختلف نمونه را ۱۰۰۰ بار تکرار نموده و با میانگین گرفتن از برآوردهای مربوط به پارامترهای مدل در ۱۰۰۰ بار تکرار، برآورد نهایی پارامترهای مدل در جدول ۱ گردآوری می‌گردد. همچنین با میانگین گرفتن از خطای استاندارد در ۱۰۰۰ بار تکرار، خطای استاندارد مربوط به پارامترهای مدل نیز به دست می‌آید. نتایج حاصل از شبیه‌سازی نشان می‌دهد که میزان اریبی پایینی برای هر کدام از برآوردکننده‌های پارامترها وجود دارد و با افزایش حجم نمونه میزان این اریبی کاهش پیدا می‌کند. همچنین مقادیر انحراف معیار با افزایش تعداد نمونه کاهش یافته‌اند. نکته حائز اهمیت آن است که مقادیر برآورد پارامترهای مدل به مقادیر واقعی آن‌ها نزدیک

جدول ۱: برآوردها و برآورد خطای استاندارد برای داده‌های شبیه‌سازی شده از مدل پواسون صفر آماسیده.

		$n = ۵۰$		$n = ۱۰۰$		$n = ۱۰۰۰$	
پارامتر	مقدار واقعی	برآورد	خطای استاندارد	برآورد	خطای استاندارد	برآورد	خطای استاندارد
$\beta_۰$	۰	۰/۱۳۳−	۰/۱۹۴	۰/۰۱۰−	۰/۰۱۷	۰/۰۰۳−	۰/۰۱۱
$\beta_۱$	۱	۰/۸۵۱	۰/۱۳۷	۰/۹۸۲	۰/۰۷۴	۰/۹۵۷	۰/۰۲۷
$\xi_۰$	۰	۰/۰۶۲−	۰/۴۳۱	۰/۰۴۴−	۰/۳۲۱	۰/۰۰۳−	۰/۰۱۶
$\xi_۱$	۱	۰/۹۵۸	۰/۵۰۲	۱/۰۱۰	۰/۳۳۹	۱/۰۸۶	۰/۱۲۳
$\sigma^۲_\eta$	۱	۰/۹۷۹	۰/۱۳۷	۱/۱۰۹	۰/۰۸۸	۱/۰۷۳	۰/۰۲۹

است و با افزایش حجم نمونه این موضوع بهتر دیده می‌شود. بنابراین می‌توان نتیجه گرفت که برآوردها سازگار هستند.

۲.۳ مدل گاما صفر آماسیده (ZIG)

فرض می‌شود متغیر پیوسته T_i دارای توزیع گاما با پارامتر $(\lambda_i^{(t)}, v)$ است. همچنین متغیر اثر تصادفی (η_i) در مدل خطی برای متغیر پاسخ T_i دارای توزیع نرمال با میانگین صفر و واریانس $\sigma^۲_{\eta_i} = ۱$ و X_i تنها متغیر تبیینی پاسخ پیوسته T_i^f بوده و نمونه‌های تصادفی مربوط به متغیر تبیینی X_i از توزیع نرمال استانداراد تولید شود، همچنین مقادیر واقعی برای بردار پارامترهای مدل به صورت

$$(\gamma_۰, \gamma_۱, \xi_۰, \xi_۱, \sigma^۲_{\eta_i})' = (۰, ۱, ۰, ۱, ۱)',$$

در نظر گرفته می‌شود. حال برای تولید نمونه تصادفی از متغیر تصادفی T_i^f به شرح زیر عمل می‌نماییم:

ابتدا متغیر تصادفی u_i از توزیع یکنواخت روی بازه (۰, ۱) تولید و سپس با در نظر گرفتن $\phi_i = ۰/۸$ و مقایسه مقادیر u_i با مقدار ثابت ۰/۸ مقادیر T_i^f به صورت زیر تولید می‌گردد:

اگر $u_i \leq ۰/۸$ ، آنگاه T_i^f مقداری برابر صفر اختیار می‌کند و اگر $u_i > ۰/۸$ ، آنگاه نمونه مورد نظر از توزیع گاما تولید می‌شود. مدل‌ها در شبیه‌سازی به صورت

$$T_i^f | \lambda_i^{(t)} \sim ZIG(\lambda_i^{(t)}, \phi_i),$$

$$\log(\lambda_i^{(t)}) = \gamma_۰ + \gamma_۱ X_i + \sigma_\eta \eta_i,$$

$$\text{logit}(\phi_i) = \xi_۰ + \xi_۱ X_i,$$

در نظر گرفته می‌شوند. در نهایت بعد از تولید نمونه تصادفی بر اساس مدل مطرح شده، با استفاده از تقریب انتگرال روش‌های $MCMC$ ، تابع لگاریتم درستنمایی مدل مورد نظر تقریب زده و با استفاده روش نیوتون رافسون برآورد پارامترهای مدل حاصل می‌شود.

شبیه‌سازی برای سه حجم نمونه ۵۰، ۱۰۰ و ۱۰۰۰ انجام می‌شود و با برازش مدل مورد نظر و به دست آوردن برآورد پارامترهای مدل این فرآیند ۱۰۰۰ مرتبه تکرار می‌گردد و سپس میانگین برآورد پارامترها و میانگین برآورد خطای استاندارد اعلام می‌گردد. نتایج حاصل از شبیه‌سازی نشان می‌دهد که میزان اریبی پایینی برای هر کدام از برآوردکننده‌های پارامترها وجود دارد و با افزایش حجم نمونه میزان این اریبی کاهش پیدا می‌کند. بنابراین می‌توان نتیجه گرفت که برآوردها سازگار هستند. لازم به ذکر است مقادیر برآورد پارامترهای مدل به مقادیر واقعی آن‌ها نزدیک است و با افزایش حجم نمونه این موضوع بهتر دیده می‌شود. بنابراین می‌توان نتیجه گرفت که برآوردها سازگار و ناریب هستند.

جدول ۲: برآوردها و برآورد خطای استاندارد برای داده‌های شبیه‌سازی شده از مدل گاما صفر آماسیده.

		$n = ۵۰$		$n = ۱۰۰$		$n = ۱۰۰۰$	
پارامتر	مقدار واقعی	برآورد	خطای استاندارد	برآورد	خطای استاندارد	برآورد	خطای استاندارد
$\beta_۰$	۰	−۰/۱۰۳	۰/۱۴۹	−۰/۰۸۱	۰/۰۳۲	۰/۰۰۳	۰/۰۰۱
$\beta_۱$	۱	۱/۳۰۰	۰/۰۸۰	۱/۳۷۰	۰/۰۰۵	۱/۳۱۰	۰/۰۰۲
$\xi_۰$	۰	۰/۰۳۱	۰/۰۶۱	۰/۰۱۱	۰/۰۴۳	۰/۰۰۳	۰/۰۰۲
$\xi_۱$	۱	۱/۳۲۱	۰/۲۱۳	۱/۰۳۱	۰/۱۱۵	۱/۰۰۹	۰/۰۱۵
$\sigma_\eta^۲$	۱	۱/۲۵۰	۰/۱۳۹	۱/۱۳۰	۰/۰۰۶	۱/۰۲۳	۰/۰۰۲

۳.۳ مدل‌های توأم پواسون-پواسون (ZIPP) و پواسون- گاما (ZIPG) صفر آماسیده

در این بخش ابتدا نتایج مربوط به شبیه‌سازی برای مدل پواسون-پواسون صفر آماسیده و پواسون-گاما صفر آماسیده ارائه می‌شود. این نتایج شامل برآوردهای پارامترهای مدل و خطای معیار آن‌ها است. شبیه‌سازی برای حالتی است که متغیر شمارشی دارای توزیع پواسون صفر آماسیده و متغیر پیوسته زمان در دو حالت زمان پیوسته و زمان گسسته تولید می‌گردد. در حالت زمان پیوسته از توزیع گاما صفر آماسیده و حالت زمان گسسته توزیع پواسون صفر آماسیده، استفاده شده است. پاسخ‌های آماسیده از این دو توزیع به این ترتیب تولید می‌شوند:
۱-متغیر زمان پیوسته با استفاده از شبیه‌سازی این مدل تولید می‌شود:

$$T_i^f|\lambda_i^{(t)} \sim ZIG(\lambda_i^{(t)}, \phi_i),$$
$$\log(\lambda_i^{(t)}) = \gamma_\circ + \gamma_\text{۱} X_i + \sigma_\eta \eta_i,$$
$$\text{logit}(\phi_i) = \xi_\circ + \xi_\text{۱} X_i$$

۲–متغیر زمان گسسته با استفاده از شبیه‌سازی این مدل تولید می‌شود:

$$T_i^f|\lambda_i^{(t)} \sim ZIP(\lambda_i^{(t)}, \phi_i),$$
$$\log(\lambda_i^{(t)}) = \gamma_\circ + \gamma_\text{۱} X_i + \sigma_\eta \eta_i.$$

در مدل‌های بالا متغیر تبیینی X_i از توزیع نرمال استاندارد تولید شده و تنها متغیر تبیینی به کار رفته در مدل است و همچنین متغیر اثر تصادفی (η_i) دارای توزیع نرمال با میانگین صفر و واریانس یک است. مقادیر اولیه مربوط به ضرایب رگرسیونی در مدل‌های بالا به صورت $(\gamma_\circ, \gamma_\text{۱}, \sigma^\text{۲}_{\eta_i})' = (\circ, \text{۱}, \text{۱})'$، است. مشابه بخش‌های قبل برای مدل‌های پواسون تک متغیره صفر آماسیده و گاما تک‌متغیره صفر آماسیده نمونه‌های تصادفی مورد نظر تولید می‌گردند. حجم نمونه برای مدل‌های شبیه‌سازی شده برابر با ۵۰، ۱۰۰ و ۱۰۰۰ خواهد بود. به طور خلاصه پاسخ‌های آمیخته توأم پواسون–گاما و پواسون- پواسون صفر آماسیده به ترتیب توسط مدل‌های زیر شبیه‌سازی می‌شوند.

۱- مدل پواسون -گاما صفر آماسیده :

$$Y_i^f|\lambda_i^{(f)} \sim ZIP(\lambda_i^{(f)}, \phi_i),$$
$$\log(\lambda_i^{(f)}) = \beta_\circ + \beta_\text{۱} X_i + \sigma_\eta \eta_i,$$
$$T_i^f|\lambda_i^{(t)} \sim ZIG(\lambda_i^{(t)}, \phi_i),$$
$$\log(\lambda_i^{(t)}) = \gamma_\circ + \gamma_\text{۱} X_i + \sigma_\eta \eta_i,$$
$$\text{logit}(\phi_i) = \xi_\circ + \xi_\text{۱} X_i.$$

۲- مدل پواسون–پواسون صفر آماسیده :

$$Y_i^f|\lambda_i^{(f)} \sim ZIP(\lambda_i^{(f)}, \phi_i),$$
$$\log(\lambda_i^{(f)}) = \beta_\circ + \beta_\text{۱} X_i + \sigma_\eta \eta_i,$$
$$T_i^f|\lambda_i^{(t)} \sim ZIP(\lambda_i^{(t)}, \phi_i),$$
$$\log(\lambda_i^{(t)}) = \gamma_\circ + \gamma_\text{۱} X_i + \sigma_\eta \eta_i,$$

$$\mathrm{logit}(\phi_i) = \xi_0 + \xi_1 X_i.$$

در جدول ۳ به ترتیب نتایج شبیه‌سازی مدل پواسون-گاما صفر آماسیده و مدل پواسون-پواسون صفر آماسیده ارائه شده‌اند. نتایج نشان می‌دهند که مدل قابلیت خوبی را در برآورد پارامترهای مدل دارد، زیرا برآوردها به مقادیر اولیه نزدیک هستند و با افزایش حجم نمونه برآوردها بهتر می‌شوند بنابراین می‌توان نتیجه گرفت که برآوردها سازگار هستند.

جدول ۳: برآوردها و برآورد خطای استاندارد برای داده‌های شبیه‌سازی شده از مدل پواسون- گاما و پواسون- پواسون صفر آماسیده.

			$n=۵۰$		$n=۱۰۰$		$n=۱۰۰۰$	
	پارامتر	مقدار واقعی	برآورد	خطای استاندارد	برآورد	خطای استاندارد	برآورد	خطای استاندارد
	β_0	۰	۰/۱۱۱	۰/۱۵۵	۰/۰۵۳	۰/۱۱۰	۰/۰۴۸	۰/۰۳۵
	β_1	۱	۰/۹۹۹	۰/۰۸۸	۰/۹۴۳	۰/۰۷۴	۰/۹۲۵	۰/۰۲۸
	γ_0	۰	۰/۵۸۱	۰/۱۳۰	۰/۳۷۸	۰/۰۹۳	۰/۴۷۶	۰/۰۲۷
ZIPG	γ_1	۱	۱/۰۴۹	۰/۰۶۱	۱/۱۶۹	۰/۰۵۲	۱/۰۴۳	۰/۰۱۹
	ξ_0	۰	−۰/۱۶۶	۰/۱۸۱	−۰/۰۸۲	۰/۱۲۶	−۰/۱۰۰	۰/۰۴۰
	ξ_1	۱	۰/۰۶۶	۰/۱۷۷	۰/۱۰۳	۰/۱۲۳	۰/۰۸۶	۰/۰۴۱
	σ_η^2	۱	۰/۹۸۸	۰/۰۷۵	۰/۰۱۰	۰/۰۶۱	۱/۰۵۰	۰/۰۱۶
	β_0	۰	−۰/۵۲۶	۰/۱۸۷	−۰/۲۲۲	۰/۱۱۲	−۰/۳۴۶	۰/۱۲۷
	β_1	۱	۱/۱۲۴	۰/۱۳۸	۰/۸۸۸	۰/۰۸۷	۱/۱۱۲	۰/۰۶۰
	γ_0	۰	−۰/۰۲۱	۲/۳۱۰	−۰/۰۳۱	۱/۶۱۲	−۰/۰۰۸	۱/۰۲۱
ZIPP	γ_1	۱	۱/۱۵۲	۰/۱۳۹	۰/۹۲۸	۰/۰۷۸	۱/۱۷۴	۰/۰۶۲
	ξ_0	۰	۰/۰۳۶	۲/۱۳۱	۰/۰۲۳	۱/۳۱۳	۰/۰۳۱	۰/۰۰۵
	ξ_1	۱	۱/۴۶۲	۰/۶۳۵	۱/۱۳۶	۰/۲۶۱	۱/۰۲۹	۰/۰۴۱
	σ_η^2	۱	۱/۵۴۰	۰/۱۵۹	۱/۰۷۱	۰/۰۳۷	۱/۰۳۹	۰/۰۵۱

۴.۳ مدل دوجمله‌ای منفی-گاما صفر آماسیده (ZINBG)

در این بخش حالتی در نظر گرفته می‌شود که متغیر شمارشی دارای توزیع دوجمله‌ای منفی صفر آماسیده و متغیر پیوسته زمان دارای توزیع گاما صفر آماسیده است. فرض می‌گردد متغیر اثر تصادفی (η_i) در مدل خطی برای متغیرهای پاسخ گسسته و پیوسته دارای توزیع نرمال استاندارد باشد. متغیر تبیینی X از توزیع نرمال استاندارد تولید می‌شود. همچنین مقادیر واقعی برای بردار پارامترهای مدل به صورت

$$(\beta_0, \beta_1, \gamma_0, \gamma_1, \xi_0, \xi_1, \sigma^2_\eta)' = (0, 1, 0, 1, 0, 1, 1)'$$

قرار داده می‌شود. مدل‌ها در شبیه‌سازی به صورت

$$\begin{aligned}
&Y_i^f|\lambda_i^{(f)} \sim ZINB(k, \lambda_i^{(f)}/k, \phi_i),\\
&\log(\lambda_i^{(f)}) = \beta_0 + \beta_1 X_i + \sigma_\eta \eta_i,\\
&T_i^f|\lambda_i^{(t)} \sim ZIG(\lambda_i^{(t)}, \phi_i),\\
&\log(\lambda_i^{(t)}) = \gamma_0 + \gamma_1 X_i + \sigma_\eta \eta_i,\\
&\text{logit}(\phi_i) = \xi_0 + \xi_1 X_i.
\end{aligned}$$

فرض می‌گردد. برای تولید نمونه تصادفی از توزیع دوجمله‌ای منفی مشابه مدل‌های قبل عمل می‌شود به این صورت که ابتدا متغیر تصادفی u_i را از توزیع یکنواخت روی بازه $(0, 1)$ تولید و سپس با در نظر گرفتن $\phi_i = 0/8$ و مقایسه مقادیر u_i با این مقدار ثابت، مقادیر Y_i^f و T_i^f به صورت زیر تولید می‌شوند: ۱- برای تولید نمونه از متغیر شمارشی دوجمله‌ای منفی، اگر متغیر تصادفی $u_i \leq 0/8$، $Y_i^f = 0$ در غیر اینصورت Y_i^f از توزیع دوجمله‌ای منفی تولید می‌شوند.

۲- برای تولید نمونه از توزیع گاما، اگر متغیر تصادفی $u_i \leq 0/8$، آنگاه $T_i^f = 0$ خواهد بود، در غیر این صورت متغیر پیوسته T_i^f از توزیع گاما تولید می‌شود.

در نهایت بعد از تولید نمونه‌های تصادفی از توزیع‌های مورد نظر به برآورد پارامترهای مدل پرداخته می‌شود. همانطور که در جدول ۴ ملاحظه شد برآورد پارامترها به مقادیر اولیه آن‌ها نزدیک هستند. همچنین دیده می‌شود که با افزایش حجم نمونه مقادیر برآورد پارامترها به مقادیر واقعی نزدیک‌تر می‌شوند و این در واقع خاصیت سازگار بودن برآوردهای ماکسیمم درستنمایی را نشان می‌دهد. مقادیر خطای استاندارد نشان می‌دهند که این مدل‌ها به خوبی برازش داده شده‌اند.

۴ مثال کاربردی

در یک ارزیابی مبتنی بر رایانه که در یک دانشگاه علمی کاربردی در هلند انجام گرفته است، از دانشجویان درخواست شده که به سؤالات (فرم‌های نظرسنجی) به صورت مجازی پاسخ دهند. این پژوهش بر روی ۶۱۰ دانشجو سال اول کارشناسی حقوق، بهداشت و مدیریت بازرگانی مورد مطالعه قرار گرفته است. یک متغیر

جدول ۴: برآوردها و برآورد خطای استاندارد برای داده‌های شبیه‌سازی شده از مدل دوجمله‌ای منفی–گاما صفر آماسیده.

		$n=۵۰$		$n=۱۰۰$		$n=۱۰۰۰$	
پارامتر	مقدار واقعی	برآورد	خطای استاندارد	برآورد	خطای استاندارد	برآورد	خطای استاندارد
$\beta_۰$	۰	۰/۰۰۳	۰/۰۱۷	۰/۰۱۹	۰/۰۱۱	۰/۰۰۲	۰/۰۰۱
$\beta_۱$	۱	۰/۹۹۱	۰/۰۱۶	۱/۰۰۲	۰/۰۰۹	۰/۹۹۹	۰/۰۰۱
$\gamma_۰$	۰	۰/۰۴۱	۱/۲۱۱	۰/۰۲۲	۱/۰۲۰	۰/۰۰۱	۰/۰۳۹
$\gamma_۱$	۱	۱/۰۰۴	۰/۰۸۵	۱/۲۰۱	۰/۰۴۲	۱/۱۹۸	۰/۰۱۷
$\xi_۰$	۰	−۰/۱۷۴	۰/۱۴۸	−۰/۱۷۷	۰/۱۰۴	−۰/۱۲۴	۰/۰۳۲
$\xi_۱$	۱	−۰/۰۱۳	۰/۱۶۷	۰/۰۶۳	۰/۰۹۲	−۰/۰۱۶	۰/۰۳۰
$\sigma_\eta^۲$	۱	۱/۰۲۴	۰/۰۱۵	۰/۹۸۸	۰/۰۱۰	۱/۰۰۲	۰/۰۰۱

مهم تعداد فرم‌های بازدید شده توسط دانشجویان است که به عنوان یک پاسخ شمارشی (Y_i^f) لحاظ گردیده است. در این مطالعه، تعداد زیادی از دانشجویان به دلایل مختلف (از جمله محدودیت زمان، انگیزه و عدم آشنایی با ارزیابی) هرگز فرم‌های نظرسنجی را بازدید نکرده‌اند. بنابراین مشاهدات مربوط به متغیر Y_i^f شامل تعداد زیادی مقدار صفر است. بنابراین توزیع پواسون برای این داده‌ها انتخاب درستی نیست. از این رو، ترجیح داده می‌شود که برای تحلیل این پاسخ شمارشی از توزیع پواسون صفر آماسیده استفاده گردد. در این مطالعه دو متغیر پاسخ شمارشی و پیوسته وجود دارد. متغیرهای شمارشی و پیوسته به ترتیب تعداد صفحات نظرسنجی بازدید شده و کل زمان پردازش بازخورد (T_i^f) را نشان می‌دهند. به پیروی از متغیر پاسخ شمارشی که تعداد زیادی مقدار صفر تولید می‌کند، متغیر پیوسته زمان نیز مقدار زیادی صفر تولید می‌کند و از این رو ترجیح داده می‌شود که برای تحلیل این پاسخ پیوسته از توزیع گاما صفر آماسیده استفاده شود. واضح است که در دو متغیر پاسخ آمیخته Y_i^f و T_i^f نوعی همبستگی وجود دارد، زیرا هر چه تعداد فرم‌های بازدید شده بیشتر باشد، زمان لازم برای پر کردن این فرم‌ها نیز بیشتر خواهد بود. در این مطالعه اطلاعاتی درباره اثرهای بازخورد بر یادگیری دانشجویان جمع‌آوری شده و هدف بررسی رفتار دانشجویان قبل و بعد از انجام آزمون مانند تعداد صفحات بازدید شده و زمان مورد انتظار برای پردازش اطلاعات بازخورد است. از جمله متغیرهایی که در این مطالعه مورد بررسی قرار گرفتند مهارت‌ها و سرعت کار دانشجویان است. این متغیرها با اندازه‌گیری عامل‌های مختلفی، مهارت و سرعت دانشجویان را در یکی از رده‌های عالی و خیلی خوب $(\theta,\zeta=۱)$، خوب و متوسط $(\theta,\zeta=۲)$ و ضعیف $(\theta,\zeta=۳)$ دسته‌بندی می‌کند.

در این قسمت نتایج حاصل ازکاربرد مدل‌هایی که در ادامه شرح داده می‌شوند، برای تحلیل داده‌های ارزیابی مبتنی بر کامپیوتر که در بخش قبل معرفی شدند، ارائه می‌گردد. متغیرهای T_i^f و Y_i^f به ترتیب متغیرهای آمیخته همبسته پیوسته و شمارشی هستند. به طوری‌که در این مطالعه متغیرهای تبیینی این دو پاسخ آمیخته θ و ζ در نظرگرفته می‌شود. با توجه به این‌که متغیر Y_i^f تعداد صفحات نظرسنجی بازدید شده است و برای این متغیر مقادیر صفر به دفعات زیادی در مشاهدات دیده شده‌اند، و به طور مشابه این پاسخ شمارشی برای متغیر زمان نیز مقادیر صفر به تعداد زیادی در مشاهدات است. مناسب‌ترین مدل برای تحلیل توأم این پاسخ‌ها مدل پواسون-گاما صفر آماسیده، پواسون-پواسون صفر آماسیده و دوجمله‌ای منفی-گاما صفر آماسیده خواهد بود. به منظور نشان دادن این امر برای تحلیل این داده‌ها پنج مدل پواسون صفر آماسیده، مدل گاما صفر آماسیده و مدل پواسون- گاما صفر آماسیده، مدل پواسون- پواسون صفر آماسیده و در نهایت مدل دوجمله‌ای منفی- گاما صفر آماسیده برازش داده شده است. برای تحلیل این پاسخ‌ها نتایج حاصل با هم مورد مقایسه قرار می‌گیرند. این مدل‌ها به ترتیب در ادامه معرفی شده‌اند.

۱- مدل پواسون صفر آماسیده:

$$
\begin{aligned}
&Y_i^f \sim ZIP(\phi_i, \lambda_i^{(f)}),\\
&\phi_i = \tfrac{\exp(z_i)}{1+exp(z_i)},\\
&z_i = \alpha_0 + \alpha_1\theta_{1i} + \alpha_2\theta_{2i} + \alpha_3\zeta_{1i} + \alpha_4\zeta_{2i},\\
&\log\lambda_i^{(f)} = \beta_{00} + \beta_{01}\theta_{1i} + \beta_{02}\theta_{2i} + \beta_{03}\zeta_{1i} + \beta_{04}\zeta_{2i} + \sigma_\eta\eta_i.
\end{aligned}
$$

۲- مدل گاما صفر آماسیده:

$$
\begin{aligned}
&Y_i^t \sim ZIG(\phi_i, \lambda_i^{(t)}),\\
&\phi_i = \tfrac{\exp(z_i)}{1+exp(z_i)},\\
&z_i = \alpha_0 + \alpha_1\theta_{1i} + \alpha_2\theta_{2i} + \alpha_3\zeta_{1i} + \alpha_4\zeta_{2i},\\
&\log\lambda_i^{(t)} = \beta_{10} + \beta_{11}\theta_{1i} + \beta_{12}\theta_{2i} + \beta_{13}\zeta_{1i} + \beta_{14}\zeta_{2i} + \sigma_\eta\eta_i.
\end{aligned}
$$

۳- مدل پواسون-گاما صفر آماسیده:

$$
\begin{aligned}
&(Y_i^f, Y_i^t) \sim ZIPG(\phi_i, \lambda_i^{(f)}, \lambda_i^{(t)}),\\
&\phi_i = \tfrac{\exp(z_i)}{1+\exp(z_i)},\\
&z_i = \alpha_0 + \alpha_1\theta_{1i} + \alpha_2\theta_{2i} + \alpha_3\zeta_{1i} + \alpha_4\zeta_{2i},\\
&\log\lambda_i^{(f)} = \beta_{00} + \beta_{01}\theta_{1i} + \beta_{02}\theta_{2i} + \beta_{03}\zeta_{1i} + \beta_{04}\zeta_{2i} + \sigma_\eta\eta_i,\\
&\log\lambda_i^{(t)} = \beta_{10} + \beta_{11}\theta_{1i} + \beta_{12}\theta_{2i} + \beta_{13}\zeta_{1i} + \beta_{14}\zeta_{2i} + \sigma_\eta\eta_i.
\end{aligned}
$$

۴- مدل پواسون-پواسون صفر آماسیده:

$$
\begin{aligned}
&(Y_i^f, Y_i^t) \sim ZIPP(\phi_i, \lambda_i^{(f)}, \lambda_i^{(t)}),\\
&\phi_i = \tfrac{\exp(z_i)}{1+\exp(z_i)},\\
&z_i = \alpha_0 + \alpha_1\theta_{1i} + \alpha_2\theta_{2i} + \alpha_3\zeta_{1i} + \alpha_4\zeta_{2i},\\
&\log\lambda_i^{(f)} = \beta_{00} + \beta_{01}\theta_{1i} + \beta_{02}\theta_{2i} + \beta_{03}\zeta_{1i} + \beta_{04}\zeta_{2i} + \sigma_\eta\eta_i,\\
&\log\lambda_i^{(t)} = \beta_{10} + \beta_{11}\theta_{1i} + \beta_{12}\theta_{2i} + \beta_{13}\zeta_{1i} + \beta_{14}\zeta_{2i} + \sigma_\eta\eta_i.
\end{aligned}
$$

۵- مدل دوجمله‌ای منفی-گاما صفر آماسیده:

$$
\begin{aligned}
&(Y_i^f, Y_i^t) \sim ZINBG(\phi_i, \lambda_i^{(f)}, \lambda_i^{(t)}),\\
&\phi_i = \tfrac{\exp(z_i)}{1+\exp(z_i)},\\
&z_i = \alpha_0 + \alpha_1\theta_{1i} + \alpha_2\theta_{2i} + \alpha_3\zeta_{1i} + \alpha_4\zeta_{2i},\\
&\log\lambda_i^{(f)} = \beta_{00} + \beta_{01}\theta_{1i} + \beta_{02}\theta_{2i} + \beta_{03}\zeta_{1i} + \beta_{04}\zeta_{2i} + \sigma_\eta\eta_i,\\
&\log\lambda_i^{(t)} = \beta_{10} + \beta_{11}\theta_{1i} + \beta_{12}\theta_{2i} + \beta_{13}\zeta_{1i} + \beta_{14}\zeta_{2i} + \sigma_\eta\eta_i.
\end{aligned}
$$

که در آن‌ها η_i دارای توزیع نرمال استاندارد است. نتایج حاصل از تحلیل پاسخ‌ها با استفاده از مدل‌های بیان شده در جدول‌های ۵ و ۶ ارائه شده‌اند. برآوردهای پارامترهای مدل پواسون- گاما اثرهای منفی برای هر دو متغیر تبیینی را نشان می‌دهند. بین توانایی وسرعت کار و تعداد کل مورد انتظار صفحات بازدید شده، ارتباط منفی وجود دارد، اگرچه دانشجوی با توانایی بالا نسبت به دانشجوی با توانایی پایین برای بازدید یک صفحه بازخورد احتمال بیشتری دارد، انتظار می‌رود که در مجموع صفحات بازخورد کمتری، بازدید شود. بر اساس

نتایج موجود به نظر می‌رسد که حدود ۴۲ درصد دانشجویان هیچ یک از صفحات بازخورد را بازدید نمی‌کنند. متغیرهای تبیینی توانایی و سرعت کار دانشجو ارتباط معنی‌داری با احتمال استفاده از بازخورد دارند، به این معنی که دانشجویانی که عملکرد خوبی در آزمون دارند به طور معنی‌داری با احتمال بالا از فرم نظرسنجی استفاده می‌کنند (یعنی حداقل یک صفحه را بازدید کردند). سرعت کار در ارتباط با احتمال استفاده از بازخورد منفی است، به این معنی که دانشجویانی که سریع کار می‌کنند نسبت به دانشجویانی که آهسته کار می‌کنند احتمالا کمتر از بازخورد استفاده می‌کنند. از مقایسه مجموع تعداد مورد انتظار صفحات بازخورد در مقابل توانایی و سرعت کار مشاهده می‌شود که تعداد صفحات بازدید شده برای دانشجویان با همان سرعت کار متفاوت است. در نهایت کل زمان پردازش بازخورد نسبت به توانایی و سرعت کار منفی است. نتایج مربوط به برآورد پارامترهای مدل‌های پواسون–پواسون صفر آماسیده نتیجه می‌شود که اثرهای متغیرهای تبیینی کاملاً مشابه مدل پواسون–گاما صفر آماسیده و دوجمله‌ای منفی–گاما صفر آماسیده است. به عبارت دیگر، تعداد کل صفحات بازدید شده با توجه به کل زمان بازخورد در مقایسه با مقدار مورد انتظار غیر شرطی کوچکتر است. به همین ترتیب، انتظار می‌رود کل زمان بازخورد به شرط تعداد کل صفحات بازدید شده نسبت به مقدار مورد انتظار غیرشرطی کوچکتر است. با توجه به مقادیر مستخرج لگاریتم تابع درستنمایی و مقادیر AIC و BIC بین مدل‌های پواسون–گاما صفر آماسیده، مدل پواسون–پواسون صفر آماسیده و مدل‌های دوجمله‌ای منفی–گاما صفر آماسیده، مدل پواسون–پواسون صفر آماسیده روی این داده‌ها به عنوان مدل مناسب انتخاب می‌شود.

بحث و نتیجه‌گیری

در این مقاله به بررسی و تحلیل پاسخ‌های آمیخته همبسته سری توانی و نرمال پرداخته شد. از آن‌جا که تحلیل این نوع پاسخ‌های در مطالعات طولی از اهمیت زیادی برخوردار هستند، می‌توان مدل‌بندی پاسخ‌ها را در حالت طولی مورد بررسی قرار داد. در این مقاله یک مدل سری توانی–گاما صفر آماسیده ارائه گردید به طوری که بررسی و تحلیل رفتار بازخورد دانشجویان در یک ارزیابی رایانه‌ای پرداخته شده است. از سویی دیگر، می‌توان این مدل‌ها را در حالتی که پاسخ‌ها دارای گم‌شدگی هستند مورد بررسی قرار داد و در نهایت مدل‌های معرفی شده را برای حالت دو متغیره تعمیم داد و نتایج مهم و پرکاربردی را به دست آورد.

جدول ۵: برآورد مدل‌های پواسون و گاما صفر آماسیده برای ارزیابی مبتنی بر رایانه.

پارامتر	ZIP		ZIG	
	برآورد	خطای استاندارد	برآورد	خطای استاندارد
α_0	۰/۵۰۵	۰/۵۳۷	−۵/۱۴$e^{-۰۶}$	۱/۵۳$e^{-۰۵}$
α_1	۰/۹۱۹	۰/۸۳۶	−۸/۳۱$e^{-۰۶}$	۶/۵۸$e^{-۰۵}$
α_2	۱/۲۵۷	۰/۶۰۲	−۲/۷۸$e^{-۰۶}$	۸/۳۲$e^{-۰۶}$
α_3	۰/۸۵۸	۰/۸۷۳	۵/۹۳$e^{-۰۶}$	۱/۷۷$e^{-۰۵}$
α_4	۰/۸۳۱	۰/۵۹۶	۱/۱۷$e^{-۰۵}$	۹/۹۶$e^{-۰۵}$
β_{00}	۰/۷۷۰	۰/۱۴۱	–	–
β_{01}	−۰/۴۸۹	۰/۱۴۸	–	–
β_{02}	−۰/۵۱۹	۰/۱۱۳	–	–
β_{03}	−۰/۳۴۱	۰/۱۴۶	–	–
β_{04}	−۰/۳۱۸	۰/۱۱۶	–	–
β_{10}	–	–	۲/۴۳	۳/۱۷$e^{-۰۲}$
β_{11}	–	–	−۴/۱۲$e^{-۰۱}$	۸/۱۰$e^{-۰۲}$
β_{12}	–	–	−۳/۳۹$e^{-۰۱}$	۵/۸۹$e^{-۰۲}$
β_{13}	–	–	−۴/۴۱$e^{-۰۱}$	۸/۶۳$e^{-۰۲}$
β_{14}	–	–	−۴/۶۱$e^{-۰۱}$	۶/۳۲$e^{-۰۲}$
σ_η^2	۱/۲۸۱	۰/۰۴۸	۱/۳۲	۲/۸۶$e^{-۰۲}$
-LogLike	۲۴۶۹/۶۱۵		۲۳۱۱/۳۴۳	
AIC	۴۹۱۶/۳		۴۶۴۴/۶۸۶	
BIC	۸۷۹۷/۳۱۴		۸۴۸۱/۷	

جدول ۶: برآورد مدل‌های پواسون–پواسون، پواسون– گاما و دوجمله‌ای منفی– گاما صفر آماسیده برای ارزیابی مبتنی بر رایانه.

	ZIPP		ZIPG		ZINBG	
پارامتر	برآورد	خطای استاندارد	برآورد	خطای استاندارد	برآورد	خطای استاندارد
α_0	۱۵/۵۸۳	۱۵/۵۰۲	۰/۲۳۸	۰/۳۷۱	−۳/۱۶$e^{-۰۶}$	۸/۸۸$e^{-۰۷}$
α_1	۲/۹۵۸	۱/۸۶۶	۰/۰۴۹	۰/۱۲۳	−۵/۶۲$e^{-۰۶}$	۷/۵۱$e^{-۱۶}$
α_2	۴/۴۲۴	۷۶/۹۶۷	۰/۰۵۵	۰/۴۱۵	−۲/۱۹$e^{-۰۶}$	۹/۶$e^{-۱۳}$
α_3	۵/۵۶۱	۵۲/۷۲۵	۰/۰۲۶	۰/۳۱۶	۸/۹۲$e^{-۰۶}$	۲/۲۶$e^{-۰۴}$
α_4	۴/۶۱۷	۸۹/۴۶۹	۰/۰۲۵	۰/۴۲۹	۱/۱۰$e^{-۰۵}$	۱/۹۸$e^{-۰۴}$
β_{00}	۰/۵۶۹	۰/۳۳۴	۰/۸۵۲	۰/۴۸۰	۱۱/۱	۵/۸۲$e^{-۰۲}$
β_{01}	−۰/۱۰۸	۰/۶۱۶	−۰/۲۵۳	۰/۴۶۱	−۳/۳۶$e^{-۰۱}$	۵/۹۶$e^{-۰۲}$
β_{02}	−۰/۳۲۵	۰/۳۶۹	−۰/۳۱۵	۰/۳۳۳	−۳/۴۶$e^{۰۱}$	۵/۲۷$e^{-۰۲}$
β_{03}	۰/۰۰۱	۰/۰۳۴	−۰/۱۶۱	۰/۵۱۷	−۲/۵۸$e^{-۰۱}$	۷/۲۳$e^{-۰۲}$
β_{04}	−۰/۳۰۹	۰/۲۹۰	−۰/۱۳۸	۰/۳۷۹	−۲/۶۴$e^{-۰۱}$	۴/۰۷۷$e^{-۰۲}$
β_{10}	۱/۲۹۱	۰/۴۱۵	۲/۳۱۳	۰/۲۳۸	۲/۷۴	۱/۰۴۱$e^{-۰۱}$
β_{11}	−۰/۱۰۳	۰/۴۵۸	−۰/۱۸۶	۰/۲۴۳	−۶/۲۱$e^{-۰۱}$	۱/۳۶$e^{-۰۱}$
β_{12}	−۰/۲۱۹	۰/۳۵۲	−۰/۲۳۳	۰/۱۷۸	−۴/۶۹$e^{-۰۱}$	۱/۱۱$e^{-۰۱}$
β_{13}	−۰/۲۲۴	۰/۳۷۱	−۰/۳۱۳	۰/۲۷۷	−۳/۲۷$e^{-۰۱}$	۱/۵۳$e^{-۰۱}$
β_{14}	−۰/۴۰۸	۰/۲۴۷	−۰/۳۰۹	۰/۱۹۴	−۳/۰۳$e^{-۰۱}$	۹/۶۰$e^{-۰۲}$
σ^2_η	۱/۶۳۰	۰/۱۲۹	۱/۰۷۶	۰/۰۶۱	۱/۰۰۷	۱/۸۵$e^{-۰۲}$
-LogLike	۴۰۵/۶۵۵۲		۸۵۴/۱۴۱۵		۷۷۳/۲۵۴	
AIC	۷۷۹/۳۱۰۴		۱۶۷۶/۲۸۳		۱۵۱۴/۵۰۸	
BIC	۱۸۷۰۸/۶۸۹۶		۱۷۸۱۱/۷۱۷		۱۷۹۷۳/۲۵۴	

تقدیر و تشکر

از داوران و ویراستار محترم مجله که با توصیه‌های بسیار مفید سبب ارتقای این مقاله و ارائه بهتر آن شده‌اند کمال تشکر را داریم.

مراجع

دهقانی، م. زاد کرمی، م. و آخوند، م. (۱۳۹۸)، مدل پواسون دو متغیره با صفر آماسیده با اثرهای تصادفی چوله نرمال و کاربرد آن، مجله علوم آماری ، مقاله آماده انتشار.

Agresti, A. (2002), *Categorical Data Analysis*, 2nd ed. Wiley, New York.

Gupta, P. L., Gupta, R. C. and Tripathi, R. C. (1996), Analysis of Zero-Adjusted Count Data, *Computational Statistics & Data Analysis*, **23**, 207-218.

Kocherlakota, S, and Kocherlakota, K. (2001), Regression in the Bivariate Poisson Distributions, *Communications in Statistics- Theory and Methods*, **30** ,815-827.

Lambert, D. (1992), Zero-Inflated Poisson Regression, With an Application to Defects in Manufacturing, *Technometrics*, **34**, 1-14.

Li, C. S., Lu, J. C., Park, J., Kim, K., Brinkley, P. A. and Peterson, J. P. (1999), Multivariate Zero-Inflated Poisson Models and Their Applications, *Technometrics*, **41**, 29-38.

Min, Y. and Agresti, A. (2005), Random Effect Models for Repeated Measures of Zero-Inflated Count Data, *Statistical Modelling*, **5**, 1-19.

Patil, M. K. and Shirke, D. T. (2011), Tests for Equality of Inflation Parameters of Two Zero-Inflated Power Series Distributions, *Communications in Statistics-Theory and Methods*, **40**, 2539-2553.

Razie, F., Samani, E. B., & Ganjali, M. (2017), Latent Variable Model for Mixed Correlated Power Series and Ordinal Longitudinal Responses with Non Ignorable Missing Values, *Communications in Statistics-Theory and Methods*, **46**, 5738-5753.

Skrondal, A. and Rabe-Hesketh, S. (2004), *Generalized Latent Variable Modeling: Multilevel, Longitudinal, and Structural Equation Models*, New York. Crc Press Company.

Tabrizi, E., Samani, E. B. and Ganjali, M. (2018), Analysis of Mixed Correlated Bivariate Zero Inflated Count and (k, l)-Inflated Beta Responses with Application to Social Network Datasets, *Communications in Statistics-Theory and Methods*, 1-31.

Van den Broek, J. (1995), A Score Test for Zero Inflation in a Poisson Distribution. *Biometrics*, **51**, 738-743.

Wang, L. (2010), IRT–ZIP Modeling for Multivariate Zero-Inflated Count Data, *Journal of Educational and Behavioral Statistics*, **35**, 671-692.

Xie, M., He, B. and Goh, T. N. (2001), Zero-Inflated Poisson Model in Statistical Process Control, *Computational Statistics & Data Analysis*, **38**, 191-201.

Yip, P. (1988), Inference About the Mean of a Poisson Distribution in The Presence of a Nuisance Parameter, *Australian & New Zealand Journal of Statistics*, **30**, 299-306.

Inflated Bivariate Random Effect Models for Mixed Power Series Normal Responses

Esmaeilzadeh, M. and Bahrami Samani, E.

Departmet of statistics, Faculty of Mathematical Science, Shahid Beheshti University, Tehran , Iran.

Abstract:

This paper will analyze inflated bivariate mixed count data. The estimations of model parameters are obtained by the maximum likelihood method. For a bivariate case which has inflation in one or two points, the new bivariate inflated power series distributions are presented. These inflated distributions are used in joint modeling of bivariate count responses. Also, to illustrate the utility of the proposed models, some simulation studies are performed. and finally, a real dataset is analyzed.

Keywords: Joint modeling, Mixed data, Inflated models, Random effects.

Mathematics Subject Classification (2010): 62J12, 62J05.