Novel Stein-type Characterizations of Bivariate Count Distributions

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH STAT

Shaochen Wang

School of Mathematics,
South China Univ. Technol., Guangzhou

Christian H. Weiß

Department of Mathematics & Statistics,
Helmut Schmidt University, Hamburg

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH
STAT

Stein Identities for Count Distributions

Introduction

Idea to uniquely characterize distribution family by certain moment identities dates back to Stein (1972, 1986).

Stein identities for various distribution families, see Betsch et al. (2022), Ebner et al. (2025), Sudheesh & Tibiletti (2012), Wang & Weiß (2023), Weiß & Aleksandrov (2022) for examples and references. In particular,

Stein identities for several **univariate count distributions**, i. e., for discrete-quantitative rv X with range in $\mathbb{N}_0 = \{0, 1, \ldots\}$.

Most relevant cases for practice are: (. . .)

(Sudheesh & Tibiletti, 2012)

Poisson distribution: $X \sim \text{Poi}(\mu)$ with $\mu > 0$ iff

$$E\big[X\,f(X)\big] \;=\; \mu\,E\big[f(X+1)\big] \qquad \ldots$$

("Stein–Chen identity" in view of Chen, 1975)

binomial distribution: $X \sim \text{Bin}(N,\pi)$ with $N \in \mathbb{N} = \{1,2,\ldots\}$ and $\pi \in (0,1)$ iff

$$(1-\pi)\,E\big[X\,f(X)\big] = \pi\,E\big[(N-X)\,f(X+1)\big] \qquad \ldots$$

negative-binomial distribution: $X \sim \text{NB}(\nu,\pi)$ with $\nu > 0$ and $\pi \in (0,1)$ iff

$$E\big[X\,f(X)\big] \;=\; (1-\pi)\,E\big[(\nu+X)\,f(X+1)\big] \qquad \ldots$$

holds for all bounded functions $f: \mathbb{N}_0 \to \mathbb{R}$.

In what follows, we derive **Stein identities for <u>bivariate</u> count distributions**, namely for

- bivariate Poisson distribution,
- bivariate binomial distribution (of type I),
- bivariate negative-binomial distribution.

We conclude with **various applications:**

- derivation of demanding moment expressions,
- bivariate goodness-of-fit tests,
- tests for symmetry (exchangeability) of bivariate distribution,
- illustrative real-world data examples.

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH STAT

Stein Characterization of Bivariate Count Distributions

General results

Let Z_0, Z_1, Z_2 be independent with $Z_i \sim \text{Poi}(\lambda_i)$, then

$$\boldsymbol{X} = (X_1, X_2)^\top = (Z_0 + Z_1,\ Z_0 + Z_2)^\top \tag{1}$$

bivariately Poisson according to $\text{BPoi}(\lambda_0; \lambda_1, \lambda_2)$.

Theorem 1: $(X_1, X_2) \sim \text{BPoi}(\lambda_0; \lambda_1, \lambda_2)$ iff

$$E[X_1\, f(X_1, X_2)] = \lambda_1\, E[f(X_1+1, X_2)] + \lambda_0\, E[f(X_1+1, X_2+1)],$$

$$E[X_2\, f(X_1, X_2)] = \lambda_2\, E[f(X_1, X_2+1)] + \lambda_0\, E[f(X_1+1, X_2+1)]$$

holds for any function $f : \mathbb{N}_0 \times \mathbb{N}_0 \to \mathbb{R}$.

Note that identities

$$E[X_1\, f(X_1, X_2)] = \lambda_1\, E[f(X_1 + 1, X_2)] + \lambda_0\, E[f(X_1 + 1, X_2 + 1)],$$

$$E[X_2\, f(X_1, X_2)] = \lambda_2\, E[f(X_1, X_2 + 1)] + \lambda_0\, E[f(X_1 + 1, X_2 + 1)]$$

include (15)–(16) in Weiß & Aleksandrov (2022) as special case.

Also note that boundary case $\lambda_0 + \lambda_2 = 0$, i. e., $X_2 \equiv 0$, leads to Stein–Chen identity for univariate Poisson distribution.

Proof: (Sketch)

Sufficiency part uses $\mathrm{BPoi}(\lambda_0; \lambda_1, \lambda_2)$'s recursive scheme

$$x\,p(x,y) = \lambda_1\,p(x-1,y) + \lambda_0\,p(x-1,y-1),$$
$$y\,p(x,y) = \lambda_2\,p(x,y-1) + \lambda_0\,p(x-1,y-1),$$

see Section 4.1.1 in Kocherlakota & Kocherlakota (2014).

Necessity part considers $f(x,y) = s^x t^y$, $|s|, |t| < 1$,

such that $G(s,t) = E[s^{X_1} t^{X_2}]$ is pgf of (X_1, X_2).

First Stein identity leads to partial differential equation

$$\frac{\partial}{\partial s} G(s,t) = \lambda_1\,G(s,t) + \lambda_0\,t\,G(s,t) = (\lambda_1 + \lambda_0\,t)\,G(s,t),$$

solved by $G(s,t) = \exp\Big\{\lambda_1(s-1) + \lambda_2(t-1) + \lambda_0(st-1)\Big\}$.

Let $\boldsymbol{Z}_k := (Z_{k1}, Z_{k2})^\top$, $k \geq 1$, be i. i. d. bivariate Bernoulli with

$$p_{ij} := P(\boldsymbol{Z}_k = (i,j)^\top), \quad i,j \in \{0,1\}.$$

Define parameters $a_1, a_2 \in (0,1)$ and $0 < a < \min\{a_1, a_2\}$ via

$$p_{11} = a, \quad p_{10} + p_{11} = a_1, \quad p_{01} + p_{11} = a_2,$$

let $\quad \phi_a := \mathrm{Corr}(Z_{k1}, Z_{k2}) = \dfrac{a - a_1 a_2}{\sqrt{a_1 a_2 (1-a_1)(1-a_2)}}.$

Then, $\quad \boldsymbol{X} = (X_1, X_2)^\top := \sum_{k=1}^N \boldsymbol{Z}_k \quad$ with $N \in \mathbb{N}$

bivariate binomial of type I, $\quad \boldsymbol{X} \sim \mathrm{BVB}(N; a_1, a_2, \phi_a)$.

According to Kocherlakota & Kocherlakota (2014), joint pgf is

$$G(z_1, z_2) := E\big[z_1^{X_1} z_2^{X_2}\big] = \big(p_{00} + p_{10} z_1 + p_{01} z_2 + p_{11} z_1 z_2\big)^N.$$

Theorem 2: $(X_1, X_2) \sim \text{BVB}(N; a_1, a_2, \phi_a)$ iff

$$p_{00}\, E[X_1 f(X_1, X_2)] + p_{01}\, E[X_1 f(X_1, X_2+1)] =$$
$$p_{10}\, E[(N-X_1)\, f(X_1+1, X_2)] + p_{11}\, E[(N-X_1)\, f(X_1+1, X_2+1)]$$

and

$$p_{00}\, E[X_2 f(X_1, X_2)] + p_{10}\, E[X_2 f(X_1+1, X_2)] =$$
$$p_{01}\, E[(N-X_2)\, f(X_1, X_2+1)] + p_{11}\, E[(N-X_2)\, f(X_1+1, X_2+1)]$$

hold for any function $f : \mathbb{N}_0 \times \mathbb{N}_0 \to \mathbb{R}$.

Stein Characterization of Bivariate Bin-Distribution

Proof analogous to Poi-case, where for sufficiency part, one first uses pgf $G(z_1, z_2)$ to derive recurrence relation

$$(x_1+1)p_{00}\, p(x_1+1, x_2) + (x_1+1)p_{01}\, p(x_1+1, x_2-1)$$
$$= (N-x_1)p_{10}\, p(x_1, x_2) + (N-x_1)p_{11}\, p(x_1, x_2-1).$$

Univariate Stein identity for boundary case $a_2 = p_{11}+p_{01} = 0$.

If $Np_{11} \to \lambda_0$, $Np_{10} \to \lambda_1$ and $Np_{01} \to \lambda_2$ for $N \to \infty$, then Theorem 2 reduces to Theorem 1 (**Poisson approximation**).

If $\nu \in (0,\infty)$, $\pi_1, \pi_2 \in (0,1)$, and $\pi_0 \in (-\pi_1\pi_2, 1)$

such that $\pi_\bullet := \sum_{i=0}^2 \pi_i < 1$ holds, then

bivariate NB-distribution $\text{BNB}(\nu; \pi_1, \pi_2, \pi_0)$ defined by pgf

$$G(z_1,z_2) = \left(\frac{1}{1-\pi_\bullet} - \frac{\pi_1}{1-\pi_\bullet} z_1 - \frac{\pi_2}{1-\pi_\bullet} z_2 - \frac{\pi_0}{1-\pi_\bullet} z_1 z_2\right)^{-\nu},$$

see Section 5 in Kocherlakota & Kocherlakota (2014).

Marginals are $X_1 \sim \text{NB}(\nu, \frac{1-\pi_\bullet}{1-\pi_2})$ and $X_2 \sim \text{NB}(\nu, \frac{1-\pi_\bullet}{1-\pi_1})$.

Theorem 3: $(X_1, X_2) \sim \mathsf{BNB}(\nu; \pi_1, \pi_2, \pi_0)$ iff

$$E[X_1\, f(X_1, X_2)] - \pi_2\, E[X_1\, f(X_1, X_2+1)] =$$

$$\pi_1\, E[(\nu + X_1)\, f(X_1+1, X_2)] + \pi_0\, E[(\nu + X_1)\, f(X_1+1, X_2+1)]$$

and

$$E[X_2 f(X_1, X_2)] - \pi_1\, E[X_2\, f(X_1+1, X_2)] =$$

$$\pi_2\, E[(\nu + X_2)\, f(X_1, X_2+1)] + \pi_0\, E[(\nu + X_2)\, f(X_1+1, X_2+1)]$$

hold for any function $f : \mathbb{N}_0 \times \mathbb{N}_0 \to \mathbb{R}$.

Proof analogous to previous cases, where for sufficiency part, one first uses pgf $G(z_1, z_2)$ to derive recurrence relation

$$x_1\, p(x_1, x_2) - \pi_2 x_1\, p(x_1, x_2 - 1)$$
$$= (\nu + x_1 - 1)\pi_1\, p(x_1 - 1, x_2) + (\nu + x_1 - 1)\pi_0\, p(x_1 - 1, x_2 - 1).$$

Univariate Stein identity for boundary case $\pi_2 = \pi_0 = 0$.

If $\nu \to \infty$ and $\nu\pi_i \to \lambda_i$ for $i = 0, 1, 2$, then Theorem 3 reduces to Theorem 1 (**Poisson limit**).

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH STAT

Stein Characterization of Bivariate Count Distributions

Applications

Moment Calculations

In analogy to Weiß & Aleksandrov (2022),

Stein identities useful for **moment computations**.

Example: Taking difference of BPoi-identities, one obtains

$$E[(X_1-X_2)\,f(X_1,X_2)] = \lambda_1\,E[f(X_1+1,X_2)]-\lambda_2\,E[f(X_1,X_2+1)].$$

Inserting $f(x,y) = x-y$, **mean squared difference** follows as

$$E[(X_1-X_2)^2] = \lambda_1+\lambda_2+(\lambda_1-\lambda_2)^2.$$

Expected absolute difference follows by $f(x,y) = \mathrm{sgn}(x-y)$.

Denote falling factorials by $x_{(k)} = x \cdots (x-k+1)$ with $x_{(0)} = 1$ and **factorial moments** by $\mu_{(r,s)} := E[(X_1)_{(r)}\,(X_2)_{(s)}]$.

Choose $f(x,y) = (x-1)_{(r-1)}\,y_{(s)}$.

From first BVB($N; a_1, a_2, \phi_a$)-identity,

$$\mu_{(r,s)} = (N-r+1)\,a_1\,\mu_{(r-1,s)} \\ - s\,a_2\,\mu_{(r,s-1)} + (N-r+1)\,s\,a\,\mu_{(r-1,s-1)}.$$

From first BNB($\nu; \pi_1, \pi_2, \pi_0$)-identity,

$$(1-\pi_\bullet)\,\mu_{(r,s)} = (\nu+r-1)\,(\pi_1+\pi_0)\,\mu_{(r-1,s)} \\ + s\,(\pi_2+\pi_0)\,\mu_{(r,s-1)} + (\nu+r-1)\,s\,\pi_0\,\mu_{(r-1,s-1)}.$$

Stein identities widely used for constructing **GoF-tests**, see Betsch et al. (2022), Weiß et al. (2023), among others.

Above BPoi-identity,

$$E[(X_1 - X_2)\, f(X_1, X_2)] = \lambda_1\, E[f(X_1+1, X_2)] - \lambda_2\, E[f(X_1, X_2+1)],$$

inspires **GoF-statistic**

$$T_{1;f} := \frac{(\mu_1 - \sqrt{\mu_1\mu_2}\,\rho)\, E[f(X_1+1, X_2)] - (\mu_2 - \sqrt{\mu_1\mu_2}\,\rho)\, E[f(X_1, X_2+1)]}{E[(X_1 - X_2)\, f(X_1, X_2)]},$$

where $E[(X_1 - X_2)\, f(X_1, X_2)] \neq 0$ required, and where deviations of $\widehat{T}_{1;f}$ from 1 indicate violation of BPoi-null.

Example: Choice $f(x,y) = x - y$ leads to kind of **bivariate dispersion index**,

$$T_{1;x-y} = \frac{\mu_1 + \mu_2 + (\mu_1 - \mu_2)^2 - 2\sqrt{\mu_1\mu_2}\,\rho}{\sigma_1^2 + \sigma_2^2 + (\mu_1 - \mu_2)^2 - 2\sqrt{\mu_1\mu_2}\,\rho},$$

and thus new competitor to

recommended dispersion-index test of Best & Rayner (1997):

$$T^* \;=\; \frac{\mu_2^2(\sigma_1^2 - \mu_1)^2 + \mu_1^2(\sigma_2^2 - \mu_2)^2 - 2\mu_1\mu_2\,(\sigma_1^2 - \mu_1)\,(\sigma_2^2 - \mu_2)\,\rho^2}{2\mu_1^2\mu_2^2(1 - \rho^4)}.$$

Comparative simulation study (10,000 replications, parametric bootstrap) uses $f_a(x, y) = x^a - y^a$ with $a \in \{1/2, 1\}$ and sample sizes $n \in \{50, 100, 200, 500\}$,

- Simulated sizes reasonably close to nominal level of 5 %.
- For BHerm- and BNB-alternatives (overdispersion), T^* most powerful for $n = 50$, but $T_{1;f_1}$ otherwise.
- For BVB-alternatives (underdispersion), $T_{1;f_1}$-test most powerful without exception, especially much more powerful than T^*-test.

Also testing for **symmetry of bivariate distribution** possible, i. e., if $p(x,y)=p(y,x)$ for all x,y.

Equivalently, (X_1,X_2) has same distribution as (X_2,X_1), i. e., X_1 and X_2 are exchangeable.

Symmetry of BPoi-distribution:

$$T_{2;f} := \Big|E[X_1\, f(X_1,X_2)] - \mu(1-\rho)\, E[f(X_1+1,X_2)]\Big| + \Big|E[X_2\, f(X_1,X_2)] - \mu(1-\rho)\, E[f(X_1,X_2+1)]\Big|$$

with alternating function f and $\mu=\frac{1}{2}(\mu_1+\mu_2)$.

Under null $(X_1,X_2)\sim \text{BPoi}(\lambda_0;\lambda,\lambda)$, it holds $T_{2;f}=0$.

Testing for Symmetry (Exchangeability)

As competitor, we consider following statistic for **symmetry of general bivariate distribution:**

$$T_{3;f} := E[f(X_1+1, X_2)] + E[f(X_1, X_2+1)]$$

with alternating function f.

Comparative simulation study (10,000 replications, parametric bootstrap) uses $f_a(x,y) = x^a - y^a$ with $a \in \{1/2, 1\}$:

- Simulated sizes reasonably close to nominal level of 5 %.
- Within BPoi, $T_{3;f}$ most powerful, closely followed by $T_{2;f_{1/2}}$.
- For non-BPoi and asymmetry, $T_{3;f}$ outperformed by any $T_{2;f}$.

Application of tests to four data sets:

- Data-1 and 2 from Crockett (1979, Tables 2–3);
- Data-3 and 4 from Best & Rayner (1997, Tables 5–6).

Descriptive statistics:

Data	n	μ_1	μ_2	σ_1^2/μ_1	σ_2^2/μ_2	ρ	μ
Data-1	100	0.940	0.640	1.415	1.216	0.276	0.790
Data-2	621	1.126	1.325	1.350	1.483	0.185	1.225
Data-3	122	1.270	0.975	1.301	1.330	0.259	1.123
Data-4	50	0.620	0.520	1.112	1.353	0.846	0.570

Test statistics (P-values in parentheses):

Data	T^*	$T_{1;f_1}$	$T_{1;f_{1/2}}$	$T_{2;f_1}$	$T_{2;f_{1/2}}$	$T_{3;f_1}$	$T_{3;f_{1/2}}$
Data-1	0.103	0.766	0.860	0.677	0.355	0.600	0.273
	(0.006)	(0.005)	(0.024)	(0.003)	(0.001)	(0.004)	(0.015)
Data-2	0.172	0.708	0.845	0.883	0.234	-0.399	-0.160
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.003)
Data-3	0.094	0.777	0.903	0.597	0.229	0.590	0.253
	(0.003)	(0.004)	(0.098)	(0.029)	(0.029)	(0.010)	(0.022)
Data-4	0.083	0.859	0.770	0.082	0.099	0.200	0.141
	(0.126)	(0.318)	(0.110)	(0.531)	(0.127)	(0.061)	(0.038)

- Data-1 to 3 not BPoi, and also not symmetric.
- Data-4 appears BPoi, and at most slightly asymmetric.

Conclusions and Future Research

- Novel Stein identities for BPoi, BVB and BNB.
- Broad potential for applications demonstrated by moment calculations, goodness-of-fit tests, symmetry tests.

Recent contributions:

- Unified approach for deriving Stein identities.
- Non-parametric implementation of $T_{3;f}$-test.

Future research:

- Optimal weight functions f for tests?
- Model diagnostics for count time series analysis based on lagged pairs (X_t, X_{t-h})?

Thank You for Your Interest!

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH
STAT

Christian H. Weiß

Department of Mathematics & Statistics

Helmut Schmidt University, Hamburg

weissc@hsu-hh.de

Literature

Wang & Weiß (2026) Unified Stein-type characterizations of bivariate count distributions with applications.
***Scandinavian Journal of Statistics*, in press (→ DOI, → arXiv).**

Best & Rayner (1997) Crockett's test of fit . . . *Biom J* **39**, 423–430.

Betsch et al. (2022) Characterizations of . . . *Elec J Stat* **16**, 1303–1329.

Chen (1975) Poisson approximation for depend. . . *Ann Prob* **3**, 534–545.

Crockett (1979) A quick test of . . . In *Interactive Statistics*, 185–191.

Ebner et al. (2025) Stein's method of . . . *Scand J Stat* **52**, 1594–1624.

Kocherlakota & Kocherlakota (2014) Bivar. . . *Wiley StatsRef* stat00972.

Stein (1972) A bound for the error . . . *Proc 6th Berkeley . . .* **2**, 583–602.

Stein (1986) *Approximate Comput. of Expectations*. IMS Lecture Notes **7**.

Sudheesh & Tibiletti (2012) Moment identity for . . . Statistics **46**, 767–775.

Wang & Weiß (2023) New character. . . *Math Comp Simul* **212**, 310–322.

Weiß & Aleksandrov (2022) Comput. (bivariate) . . . Amer Stat **76**, 10–15.

Weiß et al. (2023) Optimal Stein-type goodness. . . *Biom J* **65**, 2200073.