

## Penerapan Model Bivariat Menggunakan Copula Frank pada Jumlah gol hasil pertandingan Liga 1 Indonesia 2019

Naufal Fajar\*, Aceng Komarudin Mutaqin

Prodi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Islam Bandung, Indonesia.

\*naufalfajar1112@gmail.com, aceng.k.mutaqin@gmail.com

**Abstract.** Many studies have been conducted by researchers in various countries to process data on the number of goals scored in soccer matches for home and away teams from a soccer league. This thesis will discuss the application of the Copula Frank distribution to data on the number of goals for home and away teams from the highest caste Indonesian football league matches for 2018-2019 and to find out whether the Copula Frank distribution is a suitable opportunity distribution for modeling the case of the number of goals data. com will be used and presented as data on the results of the 2018-2019 Indonesian League 1 football match. The results of the application show that the bivariate Poisson distribution using Copula Frank is not suitable for modeling bivariate data on the number of goals for the home and away team in the Indonesian League 1 in 2018-2019.

**Keywords:** *Liga 1 Indonesia, Double Poisson Regression, Maximum Likelihood Estimation.*

**Abstrak.** Sudah banyak penelitian yang dilakukan oleh para peneliti di berbagai negara untuk mengolah data jumlah gol hasil pertandingan sepak bola untuk tim kandang dan tim tandang dari suatu liga sepak bola. Dalam Skripsi ini akan dibahas mengenai penerapan distribusi Copula Frank pada data jumlah gol tim kandang dan tim tandang hasil pertandingan Liga sepak bola Indonesia kasta tertinggi untuk tahun 2018-2019 dan untuk mengetahui apakah distribusi Copula Frank merupakan distribusi peluang yang cocok untuk memodelkan kasus data jumlah gol tersebut. Data yang akan digunakan tersebut berisi informasi jumlah gol tim kandang dan jumlah gol tim tandang pertandingan Liga 1 Indonesia tahun 2018-2019. Hasil penerapan menunjukkan bahwa distribusi Poisson bivariat menggunakan Copula Frank tidak cocok untuk memodelkan data bivariat jumlah gol tim kandang dan tim tandang Liga 1 Indonesia tahun 2018-2019.

**Kata Kunci:** *Liga 1 Indonesia, Copula Frank, Penaksir Kemungkinan Maksimum.*

## A. Pendahuluan

Sepak bola merupakan permainan beregu cabang olahraga yang terdiri atas dua kesebelasan yang saling bertanding dengan melibatkan unsur fisik, teknik, taktik, dan mental. Permainan sepak bola dimainkan dengan kurun waktu 90 menit dalam 2 babak dimana setiap babak akan berjalan selama 45 menit. Selama pergantian babak akan ada waktu 15 menit untuk melakukan istirahat dari masing-masing tim. Permainan sepak bola dipimpin oleh 4 wasit yang terdiri dari 1 wasit utama dan 2 asisten wasit dan 1 wasit cadangan. Inti dari permainan ini yaitu memasukkan bola ke gawang lawan sebanyak-banyaknya, tim yang memperoleh skor terbanyak akan menjadi pemenangnya.

Statistik sangat lekat dengan sepak bola, ilmu statistik digunakan untuk menganalisa pertandingan dengan menggunakan data-data hasil pertandingan sepak bola sebelumnya. Data tersebut adalah *shoot on target*, *shoot off target*, *foul*, dan *ball possession*. Para ilmuwan tidak hanya melihat pertandingan dan hasil suatu pertandingan saja, tetapi melakukan penelitian tentang distribusi jumlah gol dalam suatu pertandingan sepak bola. Pengaplikasian distribusi pada kasus sepak bola sudah banyak dilakukan oleh peneliti-peneliti dari berbagai negara untuk mengetahui distribusi peluang apa yang cocok untuk memodelkan jumlah gol dalam pertandingan kandang dan tandang dalam suatu Liga. Distribusi yang umum digunakan untuk memodelkan data jumlah gol kandang dan tandang diantaranya adalah distribusi Poisson dan binomial negatif. Liden (1) membahas model-model bivariat yang dapat digunakan untuk memodelkan data bivariat jumlah gol kandang dan tandang dalam suatu pertandingan sepakbola. Salah satu model bivariatnya adalah model bivariat yang menggunakan Copula diskrit. Model bivariat ini membutuhkan distribusi marginal dari data jumlah gol kandang dan tandang. Salah satu distribusi marginal tersebut adalah distribusi Poisson.

Berdasarkan latar belakang yang telah diuraikan, maka perumusan masalah dalam penelitian ini sebagai berikut: “Bagaimana menerapkan model bivariat yang menggunakan Copula diskrit dengan distribusi marginal Poisson pada data jumlah gol kandang dan tandang hasil pertandingan sepak bola Liga 1 Indonesia tahun 2018-2019?” Selanjutnya, tujuan dalam penelitian ini diuraikan dalam pokok-pokok sbb.

tujuan dari skripsi ini adalah menerapkan model bivariat yang menggunakan Copula diskrit dengan distribusi marginal Poisson pada data jumlah gol kandang dan tandang hasil pertandingan sepak bola Liga 1 Indonesia tahun 2018-2019.

## B. Metodologi Penelitian

Data yang digunakan adalah data sekunder hasil pencatatan yang diperoleh dari beberapa sumber yaitu Persatuan Sepak bola Seluruh Indonesia (PSSI) mengenai hasil pertandingan satu musim Liga 1 Indonesia tahun 2018-2019. Data yang akan digunakan tersebut berisi informasi mengenai hasil skor akhir setiap pertandingan kandang maupun tandang tim yang bertanding selama satu musim Liga 1 Indonesia tahun 2018-2019.

Pertandingan ke	Tim Kandang	Tim Tandang	Jumlah Gol Tim Kandang	Jumlah Gol Tim Tandang
1	PSS	Arema	3	1
2	Borneo	Bhayangkara	1	1
3	Bali	Persebaya	2	1
4	PSIS	Kalteng Putra	1	2
5	Persela	Madura	1	5
6	Persikabo	Badak Lampung	3	1
7	Persib	Persipura	3	0
8	PSM	Semen Padang	1	0

Pertandingan ke	Tim Kandang	Tim Tandang	Jumlah Gol Tim Kandang	Jumlah Gol Tim Tandang
9	Barito Putera	Persija	1	1
10	Bali	Bhayangkara	1	0
⋮	⋮	⋮	⋮	⋮
297	Persija	Persebaya	1	2
298	Kalteng Putra	Persija	1	3
299	Persebaya	Badak Lampung	2	1
300	Persela	Semen Padang	2	0
301	PSIS	Bhayangkara	2	3
302	Persib	PSM	5	2
303	PSS	Persikabo	5	2
304	Persipura	Borneo	2	2
305	Barito Putera	Arema	3	0
306	Bali	Madura	0	2

Berikut merupakan uraian penjelasan mengenai distribusi model bivariat yang melibatkan Copula diskrit.

Copula adalah suatu fungsi yang menggabungkan beberapa distribusi marjinal menjadi distribusi gabungan. Dalam teorema tersebut, copula digambarkan sebagai suatu fungsi yang menjarang berbagai bentuk distribusi marjinal ke suatu bentuk distribusi gabungan. Dalam ruang lingkup teori probabilitas, distribusi gabungan dapat dibentuk ke dalam struktur kebergantungan yang mempresentasikan suatu copula ke dalam distribusi-distribusi marginalnya yang dikaitkan dengan variabel acak.

Ketika dua variabel acak diskrit tidak saling bebas, fungsi distribusi kumulatif gabungannya dapat ditemukan menggunakan copula dua dimensi. Copula pada dasarnya adalah fungsi penghubung yang menghubungkan dua margin univariat bersama-sama, dan juga memungkinkan untuk parameter ketergantungan  $\theta$ . Dalam hal ini, kedua margin univariat adalah uniform standar, sehingga secara matematis fungsi copula dapat dinyatakan sebagai berikut:

$$C(v_1, v_2 | \theta) = P(U_1 \leq v_1, U_2 \leq v_2) \quad \dots (2.4)$$

Pada Persamaan di atas,  $U_1$  dan  $U_2$  adalah variabel acak yang saling bebas dan berdistribusi identik uniform standar,  $U[0,1]$ , dan  $\theta$  adalah parameter ketergantungan. Jika variabel acak  $Y_j$  memiliki fungsi distribusi kumulatif  $F_j$ , kontinu, transformasi  $F_j(Y_j)$  pasti berdistribusi uniform. Oleh karena itu, fungsi distribusi kumulatif bivariat gabungan dengan fungsi distribusi kumulatif marginal  $F_1$  dan  $F_2$  dapat dinyatakan sebagai berikut:

$$\begin{aligned} F(y_1, y_2) &= P[Y_1 \leq y_1, Y_2 \leq y_2] = P[F_1(Y_1) \leq F_1(y_1), F_2(Y_2) \\ &\leq F_2(y_2)] = P[U_1 \leq F_1(y_1), U_2 \leq F_2(y_2)] \quad \dots (2.5) \\ &= C(F_1(y_1), F_2(y_2) | \theta) \end{aligned}$$

Cara alternatif untuk menyatakan copula adalah sebagai berikut:

$$F(y_1, y_2) = F(F_1^{-1}(u_1), F_2^{-1}(u_2)) = C(u_1, u_2 | \theta) \quad \dots (2.6)$$

Ada beberapa jenis copula, yang memiliki domain berbeda dari parameter ketergantungan  $\theta$ . Untuk membangun fungsi distribusi kumulatif bivariat dari marjinal Poisson,

tipe Frank copula dipilih karena domain  $\theta \in (-\infty, \infty)$ . Sehingga Frank copula adalah pilihan yang baik dalam hal ini karena sangat umum, dan dapat dinyatakan sebagai berikut:

$$C(F_X(s), F_Y(t)) = \frac{1}{\theta} \ln \left( 1 + \frac{\exp(\theta F_X(s)) - 1)(\exp(\theta F_Y(t)) - 1)}{\exp(\theta) - 1} \right) \quad \dots (2.7)$$

Dalam hal ini,  $s, t = 0, 1, 2, \dots$  Dalam kasus distribusi Poisson, fungsi distribusi kumulatif untuk variabel acak  $X$  dan  $Y$  dapat dinyatakan sebagai berikut:

$$u_X = F_X(s) = P(X \leq s) = \sum_{m=0}^s \frac{\lambda_x^m e^{-\lambda_x}}{m!} \quad \dots (2.8)$$

$$u_Y = F_Y(t) = P(Y \leq t) = \sum_{l=0}^t \frac{\lambda_y^l e^{-\lambda_y}}{l!} \quad \dots (2.9)$$

Fungsi copula sebenarnya adalah fungsi distribusi kumulatif gabungan dan bukan fungsi massa peluang gabungan, maka peluang hasil tertentu dapat dihitung sebagai berikut:

$$P(X = 0, Y = 0) = C(F_X(0), F_Y(0)) \quad \dots (2.10)$$

$$P(X = s, Y = 0) = C(F_X(s), F_Y(0)) - C(F_X(s - 1), F_Y(0)) \quad s = 1, 2, \dots \quad \dots (2.11)$$

$$P(X = 0, Y = t) = C(F_X(0), F_Y(t)) - C(F_X(0), F_Y(t - 1)) \quad t = 1, 2, \dots \quad \dots (2.12)$$

$$P(X = s, Y = t) = C(F_X(s), F_Y(t)) - C(F_X(s - 1), F_Y(t)) - C(F_X(s), F_Y(t - 1)) + C(F_X(s - 1), F_Y(t - 1)) \quad s, t = 1, 2, \dots \quad \dots (2.13)$$

**C. Hasil Penelitian dan Pembahasan**

**Uji Chi Kuadrat**

Berikut adalah hasil uji kecocokan distribusi Poisson pada data jumlah gol tim kandang dan tim tandang menggunakan uji Chi kuadrat dengan nilai statistik uji yang diperoleh melalui bantuan perangkat lunak Microsoft Excel. Hasil uji Chi kuadrat dijelaskan pada Tabel 3 dan Tabel 4.

**Tabel 3.** Taksiran Nilai Peluang dan Nilai Harapan untuk Data Gol Tim Kandang

Kategori (i)	Jumlah Gol (x)	Frekuensi Jumlah Gol (n <sub>x</sub> )	Peluang Terjadinya Gol (p <sub>x</sub> )	Nilai Harapan Terjadinya Gol (np <sub>x</sub> )
(1)	(2)	(3)	(4)	(5)
1	0	58	0.193246	59.1333
2	1	103	0.317656	97.2028
3	2	80	0.26108	79.8905
4	3	34	0.143054	43.7744
5	4	20	0.058788	17.989
6	≥5	11	0.026176	8.0099

Berdasarkan Tabel 3 dapat diketahui bahwa nilai statistik uji Chi kuadrat adalah 3,8911. Dengan taraf nyata 5% didapat nilai kuantil distribusi Chi kuadrat dengan derajat bebas 4(6-1-

1) adalah 9,488. Diperoleh  $3,8911 < 9,488$  yang artinya nilai statistik uji lebih kecil dibandingkan dengan kuantil. Oleh karena itu, dapat disimpulkan bahwa  $H_0$  diterima, yaitu data jumlah gol tim kandang Liga 1 Indonesia pada setengah musim 2019 berasal dari populasi yang berdistribusi Poisson.

**Tabel 4.** Perhitungan Statistik Uji Chi Kuadrat Jumlah Gol Tim Tandang

Kategori (i)	Jumlah Gol (x)	Frekuensi Jumlah Gol (n <sub>x</sub> )	Peluang Terjadinya Gol (p <sub>x</sub> )	Nilai Harapan Terjadinya Gol (np <sub>x</sub> )
(1)	(2)	(3)	(4)	(5)
1	0	101	0.3346	102.3925
2	1	109	0.3663	112.0964
3	2	69	0.2005	61.3600
4	3	21	0.0732	22.3917
5	≥4	6	0.0254	7.75939
<b>Jumlah</b>		<b>306</b>	<b>1</b>	<b>306</b>

Berdasarkan Tabel 4 dapat diketahui bahwa nilai statistik uji Chi kuadrat adalah 3,1188. Dengan taraf nyata 5% didapat nilai kuantil distribusi Chi kuadrat dengan derajat bebas 3 (5-1-1) adalah 7,8147. Diperoleh  $3,1188 < 7,8147$  yang artinya nilai statistik uji lebih kecil dibandingkan dengan kuantil. Oleh karena itu, dapat disimpulkan bahwa  $H_0$  diterima, yaitu data jumlah gol tim tandang Liga 1 Indonesia pada setengah musim 2019 berasal dari populasi yang berdistribusi Poisson.

#### Penaksiran Parameter

Dalam subbab ini akan dilakukan perhitungan nilai taksiran parameter model bivariat Copula Frank yang distribusi marginalnya adalah Poisson untuk data bivariat jumlah gol tim kandang dan tim tandang Liga 1 Indonesia tahun 2018-2019 menggunakan Persamaan (2.3):

$$\hat{\lambda}_x = \frac{1}{306} \sum_{i=1}^{306} (3 + 1 + 2 \dots 2 + 3 + 0) = \frac{503}{306} = 1,6438$$

$$\hat{\lambda}_y = \frac{1}{306} \sum_{i=1}^{306} (1 + 1 + 1 \dots 2 + 0 + 2) = \frac{336}{306} = 1,0980$$

Taksiran fungsi distribusi kumulatif distribusi Poisson untuk data jumlah gol tim kandang dan jumlah data gol tim tandang disajikan dalam:

**Tabel 4.10.** Distribusi Poisson

N	Jumlah Gol Tim Kandang (x <sub>i</sub> )	Jumlah Gol Tim Tandang (y <sub>i</sub> )	(F <sub>x</sub> )	(F <sub>y</sub> )
1	3	1	0,9150	0,6997
2	1	1	0,5109	0,6997
3	2	1	0,7720	0,6997
4	1	2	0,5109	0,9008
⋮	⋮	⋮	⋮	⋮
303	5	2	0,9931	0,9008
304	2	2	0,7720	0,9008
305	3	0	0,9150	0,3335
306	0	2	0,1932	0,9008
$\hat{\lambda}_i$	<b>1,6438</b>	<b>1,0980</b>		

Dalam subbab ini akan dilakukan perhitungan nilai taksiran peluang model bivariat Copula Frank dengan distribusi marjinalnya distribusi Poisson untuk setiap pasangan data jumlah gol tim kandang dan tim tandang menggunakan Persamaan (2.7). Dengan menggunakan perangkat lunak Microsoft Excel hasil nilai taksiran peluang model Bivariat Copula Frank disajikan pada tabel 4.5. Sebagai contoh nilai taksiran peluang untuk jumlah gol tim kandang 0 dan jumlah gol tim tandang 0 adalah 0,0536, sementara nilai taksiran peluang untuk jumlah gol tim kandang 1 lalu jumlah gol tim tandang 0 yaitu 0,0339. Hasil lengkap dari taksiran nilai peluang untuk setiap jumlah gol tim kandang dan tim tandang tercantum dalam Tabel 4.5 Dengan menggunakan perangkat lunak IBM SPSS hasil nilai koefisien korelasi  $\tau$  Kendall yaitu 0,053. Syntax pada nilai taksiran tersebut ada di Lampiran 2.

Langkah selanjutnya menghitung nilai taksiran parameter kebergantungan  $\theta$  sebagai solusi dari Persamaan (2.10) sampai (2.13). Dengan menggunakan perangkat lunak Matlab hasil nilai taksiran Parameter kebergantungan  $\theta$  yaitu 1,6098. Syntax pada nilai taksiran tersebut ada di Lampiran 3.

**Tabel 4.11** Taksiran Nilai Peluang

		Jumlah Gol Tim Tandang							
		0	1	2	3	4	5	6	7
jumlah gol Tim Kandang	0	0,0536	0,0921	0,0694	0,0292	0,0084	0,0019	0,000341	0,000005763
	1	0,0339	0,0485	0,0301	0,0113	0,0031	0,0007	0,0001	0,000019639
	2	0,0409	0,0508	0,0278	0,0098	0,0026	0,0006	0,0001	0,000016523
	3	0,0284	0,0313	0,0155	0,0052	0,0014	0,0003	0,000055	0,0000086
	4	0,0129	0,0133	0,0062	0,0020	0,0005	0,0001	0,000021	0,000003346
	5	0,0044	0,0044	0,0020	0,0006	0,0001	0,000037	0,000007	0,00000106
	6	0,0019	0,0012	0,0005	0,0002	0,000046	0,00001	0,000002	0,000000283
	7	0,0003	0,0003	0,000123	0,00004	0,00001	0,000002	0,000000414	0,000000065

Selanjutnya akan dilakukan perhitungan nilai harapan model bivariat Copula Frank dengan distribusi marjinalnya distribusi Poisson untuk setiap pasangan data jumlah gol tim kandang dan tim tandang. Dengan menggunakan perangkat lunak Microsoft Excel hasil nilai harapan disajikan pada tabel 4.12. Sebagai contoh nilai harapan untuk jumlah gol tim kandang 0 dan jumlah gol tim tandang 0 adalah 16, sementara nilai harapan jumlah gol tim kandang 1 dan jumlah gol tim tandang 0 yaitu 10. Hasil lengkap dari nilai harapan untuk setiap jumlah gol tim kandang dan jumlah gol tim tandang tercantum dalam:

**Tabel 4.12.** Nilai Harapan Jumlah Gol

		nilai harapan jumlah gol kandang							
		0	1	2	3	4	5	6	7
nilai harapan jumlah gol tandang	0	16	28	21	9	3	1	0	0
	1	10	15	9	3	1	0	0	0
	2	13	16	9	3	1	0	0	0
	3	9	10	5	2	0	0	0	0
	4	4	4	2	1	0	0	0	0
	5	1	1	1	0	0	0	0	0
	6	0	0	0	0	0	0	0	0
	7	0	0	0	0	0	0	0	0

Terlihat bahwa nilai harapan jumlah gol tim kandang dan jumlah gol tim tandang tidak dekat dengan nilai sebenarnya. Jadi diduga bahwa model bivariat Copula Frank dengan

distribusi marginal Poisson tidak cocok untuk data bivariat jumlah gol tim kandang dan jumlah gol tim tandang hasil pertandingan Liga 1 Indonesia tahun 2018-2019

#### D. Kesimpulan

Berdasarkan hasil analisis penerapan model bivariat Copula Frank dengan distribusi marginal Poisson pada data bivariat jumlah gol tim kandang dan tim tandang hasil pertandingan sepak bola Liga 1 Indonesia tahun 2018-2019 dapat disimpulkan bahwa modelnya tidak cocok.

#### Acknowledge

Penelitian ini dipersembahkan untuk semua pihak yang telah membantu penulis. Terima kasih untuk Allah SWT. yang telah memberikan rahmat serta hidayah-Nya, kedua orang tua yang tak henti-hentinya memanjatkan do'a dan memberi dukungan, Bapak Dr. Aceng Komarudin Mutaqin, M.T., M.Si., yang telah membimbing penulis dengan sabar dan tulus, seluruh dosen dan laboran Statistika UNISBA yang telah memberikan ilmu dan dedikasinya, serta teman-teman yang selalu memberi semangat kepada penulis.

#### Daftar Pustaka

- [1] Fitriawati Andi & Febrianti Werry,. “*Teknik Mengkontruksi Distribusi Copula Clayton Pada Data Marginal Diskrit Dengan Implikasi Bergantungan*”.(2020). Skripsi tidak dipublikasikan. Sumatera: Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Sumatera.
- [2] Hartadinata Yudi, Wahjono.T.D & Ohyver Margaretha. “*Analisis Peramalan Hasil Pertandingan Sepakbola EPL Dengan Distribusi Poisson Berbasis Komputer*”. (2011). Skripsi tidak dipublikasikan. Jakarta: Jurusan Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam, BINUS University.
- [3] Ibrahim.G.S. “*Perbandingan Penerapan Distribusi Poisson, Binomial Negatif dan Zero Inflated Poisson Pada Data Jumlah Gol Hasil Pertandingan Sepakbola Indonesia*” .(2021). Skripsi tidak dipublikasikan. Bandung: Jurusan Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Islam Bandung.
- [4] Lidén, Joel. “*Bivariate Models to Predict Football Results.*” (2016). Desertasi publikasikan. Seattle:Departmen of Mathematics, Uppsala Universitet.
- [5] Nahliyani. “*Pemilihan Model Copula Menggunakan Distribusi Copula Empirik*”.(2018). Skripsi tidak dipublikasikan.. Makassar: Jurusan Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Hassanudin.
- [6] Trivedi.P.K & Zimmer.M.D “*Copula Modelling for Practioners*”. (2005). Jurnal dipublikasikan. Bloomington: Departmen of Economics, Western Kentucky University.
- [7] Wachidah Lisnur. “*Uji Kecocokan Chi Kuadrat Untuk Distribusi Poisson*”. (2009). Jurnal dipublikasikan. Yogyakarta: Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Yogyakarta.