Jurnal Matematika, Komputasi dan Statistika
Volume 4 Nomor 1, pp. 491-497, Januari-April 2024,
ISSN: 2503 - 2984

# UJI *GOODNESS OF FIT* UNTUK DISTRIBUSI GEOMETRIK MENGGUNAKAN UJI STATISTIK KOLMOGOROV-SMIRNOV

**Reski Amelia**[1)]
[1)]Program Studi Matematika, Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Halu Oleo, Kendari, Indonesia
Email: reskiameliaa05@gmail.com

**Wayan Somayas**[1,a)], **Alfian** [1,b)] dan **Ruslan**[2,c)]
[1)]Program Studi Statistika, Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Halu Oleo, Kendari, Indonesia
[2)]Program Studi Statistika, Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Halu Oleo, Kendari, Indonesia
Email: [a)] wayan.somayasa@uho.ac.id, [b)] alfianmath03@uho.ac.id dan [c)]ruslan@uho.ac.id

**ABSTRAK**

Penelitian ini bertujuan untuk mengevaluasi kecocokan distribusi geometrik dengan data lemparan bola basket. Estimasi parameter ($P$) distribusi geometrik dengan metode *MLE* diperoleh dengan persamaan $p = \frac{1}{\bar{x}}$. Ketakbiasan penduga *MLE* diuji dengan membandingakn nilai variansi dan nilai *BBCR*, diperoleh nilai var $\neq BBCR$ sehingga $\frac{1}{\bar{x}}$ bukan merupakn *UMVUE*. Hasil analisis statistik menggunakan uji Kolmogorov-Smirnov menunjukkan bahwa data tersebut mengikuti distribusi geometrik dengan parameter $p = \frac{1}{8{,}52}$ . Statistik uji Kolmogorov-Smirnov = 4,0135 lebih besar dari nilai tabel kuantil-kuantil = 0,3667 pada tingkat signifikansi α = 0,05 untuk n = 25. Oleh karena itu, hipotesis nol yang menyatakan bahwa data mengikuti distribusi geometrik diterima, menunjukkan kesesuaian yang signifikan antara distribusi geometrik yang diasumsikan dan distribusi sebenarnya dari data lemparan bola basket.

**Kata Kunci:** Uji Kolmogorov-Smirnov, Distribusi Geometrik

***ABSTRACT***

*This study aims to evaluate the suitability of the geometric distribution to basketball throwing data. The parameter estimation (P) of the geometric distribution using the MLE method is obtained by the equation $p = \frac{1}{\bar{x}}$. The unbiasedness of the MLE estimator was tested by comparing the variance value and the BBCR value, we obtained a value of var ≠ BBCR so that $p = \frac{1}{\bar{x}}$. is not a UMVUE. The results of statistical analysis using the Kolmogorov-Smirnov test show that the data follows a geometric distribution with parameter $p = \frac{1}{8{,}52}$. The Kolmogorov-Smirnov test statistic = 4.0135 is greater than the table value of quantile-quantile = 0.3667 at the significance level α = 0.05 for n = 25. Therefore, the null hypothesis which states that the data follows a geometric distribution is accepted, indicating significant agreement between the assumed geometric distribution and the actual distribution of basketball throwing data.*

***Keywords:*** *Kolmogorov-Smirnov Test, Geometric Distribution*

## 1. Pendahuluan

Uji *goodness of fit* diperkenalkan pertama kali oleh Karl Pearson pada tahun 1902. Uji *goodness of fit* merupakan metode yang digunakan untuk mengetahui model distribusi peluang suatu populasi dari suatu sampel yang diambil. Dengan kata lain, uji ini digunakan untuk menguji hipotesis bahwa populasi dari suatu sampel memiliki fungsi distribusi tertentu[1].

Salah satu metode yang sering digunakan dalam melakukan prosedur uji *goodness of fit* adalah berdasarkan fungsi distribusi empiris. Metode pengujian ini berdasarkan pada jarak antara fungsi distribusi empiris $\hat{F}_n(\cdot)$ dan $F_0(\cdot)$. Uji *goodness of fit* berdasarkan fungsi distribusi empiris yang populer, yaitu uji Kolmogorov-Smirnov, uji Cramér-Von

Mises, dan uji Anderson-Darling. Studi tentang pengembangan prosedur uji *goodness of fit* berdasarkan CDF (*Cumulative Distribution Function*) empiris telah banyak dilakukan baik dari aspek matematika maupun terapannya [2].

Langkah-langkah uji *goodness of fit* dapat digunakan dalam pengujian hipotesis statistik, misalnya untuk menguji normalitas residu, untuk menguji apakah sampel diambil dari distribusi yang identik, atau apakah frekuensi hasil mengikuti distribusi yang diberikan. Uji normalitas adalah suatu prosedur uji *goodness of fit* yang lebih khusus dimana melibatkan pengujian hipotesis apakah data dari distribusi peluang tertentu mengikuti distribusi normal [3].

Dalam penerapan distribusi geometrik dapat menggunakan data lemparan bola basket ke ring. Pada setiap percobaan dilakukan sampai sukses. Berdasarkan hal tersebut maka dalam proposal ini akan ditentukan distribusi sampling peluang dari bola yang akan dimasukan ke ring, setiap populasi dengan menggunakan uji goodness of fit untuk distribusi geometrik menggunakan uji statistik Kolmogorov-Smirnov.

Dalam penerapanya uji kolmogorov-smirnov banyak digunakan untuk menguji apakah data sampel berasal dari distribusi normal atau tidak, dalam penelitian ini saya tertarik untuk menerapkan atau menguji data sampel yang mengikuti karakteristik distribusi geometrik diuji menggunakan statistik kolmogorov-smirnov.

## 2. Kajian Pustaka

### 2.1. Distribsui Geometrik

Distribusi geometrik adalah kasus khusus dari distribusi binomial negatif untuk $k = 1$, yaitu distribusi peluang banyaknya percobaan yang diperlukan untuk mendapatkan sukses pertama. Distribusi ini berpangkal pada percobaan Bernoulli dimana hanya terdapat dua kemungkinan yaitu sukses atau gagal, yang diulang berkali-kali sampai mendapatkan sukses pertama. Dimana setiap percobaan tidak akan berpengaruh pada percobaan selanjutnya.

**Definisi 1 [4]** Variabel acak yang didefinisikan sebagai banyaknya trial yang dilakukan sampai diperoleh sukses yang pertama disebut berdistribusi geometrik. Suatu variabel acak $X$ dikatakan berdistibusi geometrik dengan peluang sukses $p$, ditulis $X \sim GEO\ (p)$, jika $X$ mempunyai persamaan berbentuk:

$$p(x) = g(x; p) = pq^{x-1}$$

dimana $x = 1, 2, 3, \dots;$ $p$ dan $q$ adalah parameter sukses dan gagal.

Fungsi $f(x)$ disebut fungsi kepadatan peluang (distribusi geometrik) dari peubah acak $X$ diskrit, jika memenuhi kedua syarat fungsi peluang:

1. Untuk syarat yang pertama $0 \leq f(x) \leq 1$
2. $\sum_x f(x) = 1$

nilai harapan dari $X, E(X)$ diberikan oleh:

$$E(X) = \sum_{x=1}^{\infty} x.pq^{x-1}$$

$$E(X) = \frac{1}{p}$$

$$Var(X) = E(X^2) - [E(X)]^2 = \frac{1+q}{p^2} - \frac{1}{p^2} = \frac{q}{p^2} = \frac{1-p}{p^2}$$

### 2.2. Statistik dan Distribusi Sampling

Statistik bertujuan untuk mendapatkan gambaran secara terperinci mengenai karakteristik data sehingga nantinya dapat ditarik kesimpulan. Alat utama yang digunakan adalah apa yang disebut sebagai statistik.

**Definisi 2 [4]** Sebuah fungsi dari variabel acak teramati, $T = t(X_1, \dots, X_n)$ yang tidak tergantung pada suatu parameter yang tidak diketahui yang disebut statistik. Selanjutnya model distribusi peluang dari suatu statistik disebut distribusi sampling.

**Definisi 3 [5]** Suatu himpunan variabel acak $\{X_1, \dots, X_n\}$ dikatakan sebagai sampel acak berukuran $n$ dari suatu populasi $X$, Jika dan hanya jika $X_1, \dots, X_n$ saling bebas (mutually independent) dan model berdistribusi marginal dari $X_i$ identik dengan model distribusi dari $X$. Dalam beberapa literatur $\{X_1, \dots, X_n\}$ sering juga disebut duplikat (*copy*) atau perulangan yang saling bebas dari $X$.

### 2.3. Fungsi Distribusi Empiris

Fungsi distribusi empiris merupakan fungsi distribusi yang dikontruksikan dari sampel acak dari populasi.

**Definisi 4 [5]** Misalkan $X_1, \dots, X_n$ adalah sampel acak berukuran $n$ dari suatu populasi $X$. Fungsi distribusi empiris dari sampel acak tersebut adalah fungsi yang dinotasikan dengan $\widehat{F}: \mathbb{R} \rightarrow [0,1]$ yang didefinisikan sebagai

$$\widehat{F_n}(x) := \frac{1}{n} |\{X_i \leq x_i, i = 1, \dots, n\}|, \forall x \in \mathbb{R}$$

dimana tanda mutlak $|A|$ menyatakan banyaknya anggota dari himpunan $A$. Untuk setiap $x \in \mathbb{R}$, fungsi distribusi empiris dari sampel menyatakan proporsi sampel yang lebih kecil atau sama dengan $x$. Dari definisi terlihat bahwa fungsi distribusi empiris dari suatu sampel acak merupakan fungsi acak (random) yang berarti bahwa untuk setiap $x \in \mathbb{R}, \widehat{F_n}(x)$ mengikuti suatu model distribusi tertentu.

### 2.4. Estimasi Maksimum Likelihood

*Maximum likelihood estimation* atau dikenal juga sebagai metode estimasi *likelihood* maksimum adalah salah satu metode yang paling sering digunakan untuk mencari nilai estimasi parameter dari suatu distribusi.

**Definisi 5 [5]** Misalkan $X_1, \cdots, X_n$ merupakan sampel acak berukuran $n$ dari populasi dengan PDF $f_X(x; \theta_1, .., \theta_k)$, untuk $(\theta_1, .., \theta_k) \in \Theta$. Misalkan $x_1, .., x_k$ merupakan data atau suatu realisasi dari

$X_1,..,X_n$. Fungsi *likehood* untuk $(\theta_1,..,\theta_k)$ adalah fungsi $L:\Theta \to \mathbb{R}_{\geq 0}$, sedemikian hingga

$$L(\theta_1,..,\theta_k) = f_{X_1},\dots,X_n((x_1,..,x_n;\theta_1,..,\theta_k) = \prod_{i=1}^{n} f_{X_i}(x_i;\theta_1,..,\theta_k)$$

Penduga MLE untuk $(\theta_1,..,\theta_k)$ adalah statistik $(\hat{\theta}_i(X_1,..,X_n), i = 1,\dots,k,$ sedemikian hingga $L\left(\hat{\theta}_i(x_1,..,x_n),\dots,\hat{\theta}_k(x_i,..,x_n)\right) = \max_{(\theta_1,..,\theta_k)\in\Theta} L(\theta_1,\dots,\theta_k)$.

Dari **definisi 5** jelas bahwa permasalahan menentukan MLE merupakan permasalahan optimasi. Nilai- nilai $(\hat{\theta}_1,..,\hat{\theta}_k)$ memberikan nilai maksimum global dari $L(\theta_1,..,\theta_k)$. Karena nilai-nilai dari $(\theta_1,..,\theta_k)$ yang maksimumkan $L(\theta_1,..,\theta_k)$ juga sekaligus memaksimumkan fungsi log-*likelihood*, yaitu $\ln L(\theta_1,..,\theta_k)$, maka untuk memudahkan perhitungan, cukup diperhatikan $\ln L(\theta_1,..,\theta_k)$ saja.

Adapun langkah yang digunakan untuk mendapatkan estimasi parameter dengan metode estimasi maksimum *likelihood* adalah sebagai berikut

1) Mendefiniskan fungsi *likelihood* $L(\theta_1,\theta_2,\dots,\theta_k)$
2) Mengoperasikan fungsi *likelihood* dengan logaritma natural $\ln L(\theta_1,\theta_2,\dots,\theta_k)$
3) Mendeferensialkan $\ln L(\theta_1,\theta_2,\dots,\theta_k)$ terhadap $\theta_1,\theta_2,\dots,\theta_k$ dan menyamakan derivatifnya dengan nol.

Menyelesaikan derivatif tersebut dalam $\theta_1,\theta_2,\dots,\theta_k$ akan diperoleh $\hat{\theta}_1,\dots,\hat{\theta}_k$. [5]

### 2.5. Uji Hipotesis

Uji hipotesis dikembangkan sebagai suatu usaha untuk memberikan jawaban atas pertanyaan-pertanyaan. Hipotesis adalah dugaan atau pernyataan sementara yang digunakan untuk menyelesaikan suatu permasalahkan dalam penelitian yang kebenarannya harus diuji secara empiris.

**Definisi 6 [5]** Misalkan $X$ adalah populasi kontinu atau diskrit dengan PDF $f_X(x;\theta), \theta \in \Theta \subseteq \mathbb{R}$. Hipotesis statistik adalah pernyataan tentang distribusi peluang dari $X$. Dalam kasus parametrik, hipotesis statistik adalah penyataan tentang parameter $\theta$. Tujuan dari uji hipotesis adalah untuk menjawab apakah pernyataan tersebut benar atau tidak. Suatu hipotesis dikatakan sederhana, jika hipotesis tersebut tidak menyebutkan secara jelas model distribusi peluang dari populasi, maka disebut hipotesis majemuk.

### 2.6. Uji Kolmogorov-Smirnov

Konsep dasar dari uji *goodness of fit* Kolmogorov-Smirnov adalah dengan membandingkan fungsi distribusi empiris sampel dengan fungsi distribusi dari populasi.

Diberikan data $x_1,\dots,x_n$ yang dianggap realisasi dari suatu sampel acak $X_1,\dots,X_n$ yang diambil dari suatu populasi $X$ dengan $F_n$ adalah fungsi distribusi kumulatif (CDF) yang tidak diketahui untuk populasi $X$ dan $F_0(x,\theta)$ adalah suatu fungsi distribusi kumulatif (CDF) yang diasumsikan. Misalkan $F_0(x,\theta)$ adalah CDF gamma berbobot, normal, eksponensial, Poisson, dan lain-lain.

Cocok atau tidaknya $F_0(x,\theta)$ sebagai CDF populasi dapat ditunjukan melalui uji terhadap hipotesis berikut:

$$\{H_0: F(x) = F_0(x), \forall\, x \in \mathbb{R}\ \ vs\ \ H_1: F(x) \neq F_0(x) \text{ untuk suatu x} \in \mathbb{R}\}$$

Dimana $F_x$ adalah CDF dari populasi yang tidak diketahui.

Statistik uji Kolmogorov-Smirnov yang tepat untuk menyelesaikan hipotesis tersebut adalah sebagai berikut:

$$D_n = \sup_{|x|<\infty}\left|\hat{F}_n(x) - F_0(x;\theta)\right|$$

$\hat{F}_n(x)$ adalah fungsi distribusi kumulatif yang tidak diketahui dari populasi $x, F_0(x;\theta)$ adalah suatu distibusi yang diasumsikan, $F_0$ adalah CDF gamma, berbobot normal, eksponensial, poison dan lain-lain, $F_0(x)$ adalah CDF dari distribusi geometrik. Berdasarkan statistik ini maka $H_0$ ditolak jika nilai $D$ sama atau lebih besar dari nilai $D_n$, $\alpha$ pada tabel Kolmogorov Smirnov $(D_n \geq D_n, \alpha)$ [6].

## 3. Hasil dan Pembahasan

### 3.1. Estimasi Parameter Distribusi Geometrik

$$\begin{aligned} L(p) &= \prod_{i=1}^{n} f(x_i;p) \\ &= \prod_{i=1}^{n} p(1-p)^{x_i}, \qquad x_i = 1,2,3,\dots n \\ &= p^n(1-p)^{\Sigma_{x_i}-n} \end{aligned}$$

Maka logaritma fungsi *likelihood*-nya adalah

$$\begin{aligned} lnL(p) &= ln\,(p^n(1-p)^{\Sigma_{x_i}-n}) \\ &= n.lnp + (\Sigma_{x_i} - n).ln\,(1-p) \end{aligned}$$

Penduga maksimum *likelihood* untuk $p$ diperoleh dari penyelesaian $\frac{\partial \ln L(p)}{\partial p} = 0$.

$$\frac{\partial \ln L(p)}{\partial p} = 0 \Leftrightarrow \frac{\partial \ln L\left(nlnp + (\Sigma_{x_i}-n)ln\,(1-p)\right)}{\partial p} = 0$$

$$\Leftrightarrow \frac{n}{p} + \frac{\Sigma_{x_i}-n}{1-p}.(-1) = 0$$

$$\Leftrightarrow p = \frac{1}{\bar{x}}$$

Karena $\hat{p} = \frac{1}{\bar{x}}$ merupakan satu-satunya penyelesaian, maka $\frac{1}{\bar{x}}$ adalah penduga MLE untuk $p$ jika $\frac{d^2 \ln L(p)}{dp^2} < 0$ yang dikenal pada uji turunan kedua pada titik $p = \frac{1}{\bar{x}}$. Hasil uji turunan kedua adalah sebagai berikut

$$\left.\frac{\partial^2 \ln L(p)}{\partial p^2}\right|_{p=\frac{1}{\bar{x}}} = \frac{\partial}{\partial p}\left(\frac{n}{p} - \frac{\Sigma_{x_i}-n}{1-p}\right)$$

$$= -n\,.\,p^{-2} - (-1)(\Sigma xi - n)(1-p)^{-2}(-1)\Big|_{p=\frac{1}{\bar{x}}}$$

$$= -\frac{n}{p^2} - \left.\frac{\Sigma_{x_i}-n}{(1-p)^2}\right|_{p=\frac{1}{\bar{x}}}$$

$= -n\bar{x}^2\left(1+\frac{1}{(\bar{x}-1)}\right); \bar{x} \neq 1$

Karena $\hat{p} = \frac{1}{\bar{x}}$, dimana $p$ merupakan probabilitas dalam hal ini $\bar{x} > 1$, sehingga nilai $\frac{1}{(\bar{x}-1)} > 0$, maka dapat disimpulkan bahwa $-n\bar{x}^2\left(1+\frac{1}{(\bar{x}-1)}\right) < 0$. Jadi MLE untuk $p$ adalah $\hat{p} = \frac{1}{\bar{x}}$.

### 3.2. Sifat Ketakbiasan Penduga *Likelihood* Maksimum

**Definisi 7 [5]** Misalkan $X_1, \dots, X_n$ merupakan sampel acak dari suatu populasi yang mempunyai pdf $f(x;\theta)$, $\theta \in \Theta$. Jika $T := t(X_1, \dots, X_n)$ merupakan penduga tak bias untuk $\tau(\theta)$, jika $\tau'(\theta) := \frac{\partial\tau(\theta)}{\partial\theta}$ ada. Maka berlaku

$$Var(T) \geq \frac{\left(\frac{\partial\tau(\theta)}{\partial\theta}\right)^2}{nE\left(\frac{\partial \ln f(X;\theta)}{\partial\theta}\right)^2}$$

besaran $\frac{\left(\frac{\partial\tau(\theta)}{\partial\theta}\right)^2}{nE\left(\frac{\partial \ln f(X;\theta)}{\partial\theta}\right)^2}$ disebut batas bawah Cramer – Rao

Jika suatu penduga T bersifat tak bias terhadap $\tau(\theta)$ memenuhi var (T) = batas bawah Cramer-Rao, maka T merupakan penduga UMVUE (*Unifromly Minimum Variance Unbiased Estimator*) untuk $\tau(\theta)$. Penduga UMVUE merupakan penduga terbaik.

$$\begin{aligned} \text{Var}(\bar{X}) &= \text{var}\left(\frac{1}{n}\sum_{i=1}^n X_i\right) \\ &= \frac{1}{n^2}\sum_{i=1}^{n} var(X_i) \\ &= \frac{1}{n^2} n \frac{1-p}{p^2} \\ &= \frac{1-p}{np^2} \end{aligned}$$

$$BBCR = \frac{\left(\frac{\partial\tau(p)}{\partial p}\right)^2}{nE\left(\frac{\partial \ln f(X;p)}{\partial p}\right)^2}$$

a. $\left(\frac{\partial\tau(p)}{\partial p}\right)^2 = \left(\frac{\partial\tau(p)}{\partial p}\right)^2$
$= \left(\frac{\partial(p)}{\partial p}\right)^2$
$= 1$

b. $$\begin{aligned} nE\left(\frac{\partial(\ln f(X;p))}{\partial p}\right)^2 &= nE\left(\frac{\partial(\ln(p(1-p)^{x-1}))}{\partial p}\right)^2 \\ &= nE\left(\frac{\partial(\ln(p)+(X-1)\ln(1-p))}{\partial p}\right)^2 \\ &= nE\left(\frac{(X-1)}{(1-p)} - \frac{1}{p}\right)^2 \\ &= nE\left(\frac{p(X-1)}{p(1-p)} - \frac{(1-p)}{p(1-p)}\right)^2 \\ &= nE\left(\frac{pX-p-1+p}{p(1-p)}\right)^2 \\ &= nE\left(\frac{pX-1}{p(1-p)}\right)^2 \\ &= n\frac{1}{(p(1-p))^2}E\left(p\left(X-\frac{1}{p}\right)\right)^2 \\ &= \frac{np^2}{(p(1-p))^2}E\left(X-\frac{1}{p}\right)^2 \\ &= \frac{np^2}{(p(1-p))^2}Var(X) \\ &= \frac{np^2}{(p(1-p))^2}\frac{1-p}{p^2} \\ &= \frac{n(1-p)}{p^2(1-p)^2} \\ &= \frac{n}{p^2(1-p)} \end{aligned}$$

Sehingga

$$\begin{aligned} BBCR &= \frac{\left(\frac{\partial\tau(p)}{\partial p}\right)^2}{nE\left(\frac{\partial \ln f(X;p)}{\partial p}\right)^2} \\ &= \frac{1}{\left(\frac{n}{p^2(1-p)}\right)} \\ &= \frac{p^2(1-p)}{n} \end{aligned}$$

Karena Var $(\bar{X}) = \frac{1-p}{np^2}$ dan BBCR $= \frac{p^2(1-p)}{n}$ maka Var $(\bar{X}) \neq$ BBCR jadi $\bar{X}$ bukan merupakan UMVUE.

### 3.3. MSE *(Mean Square Error)*

**Definisi 8 [7]** misalkan T merupakan sembarang pendug untuk $\tau(\theta)$. Bias dari T teehadap $\tau(\theta)$ dinotisakn sebagai b(T) adalah

$$b(T) = E(T) - \tau(\theta)$$

Sedangkan, MSE dari T adalah

$$\text{MSE}(T) = E(T-\tau(\theta))^2$$

**Teorema 1 [7]** Jika T merupakan suatu penduga untuk $\tau(\theta)$, maka

$$\text{MSE}(T) = Var(T) + (b(T))^2$$

Maka

$$\begin{aligned} var(\bar{X}) &= var\left(\frac{1}{n}\sum_{i=1}^n X_i\right) \\ &= \frac{1}{n^2}\sum_{i=1}^n var(X_i) \\ &= \frac{1}{n^2} n \frac{1-p}{p^2} \\ &= \frac{(1-p)}{np^2} \end{aligned}$$

Dan

$$\begin{aligned} b(T) &= E(T-\tau(p))^2 \\ &= \left(\left(\frac{1}{n}\sum_{i=1}^n \frac{1}{p}\right) - \frac{1}{p}\right)^2 \\ &= \left(\left(\frac{1}{n} n \frac{1}{p}\right) - \frac{1}{p}\right)^2 \\ &= \left(\frac{1}{p} - \frac{1}{p}\right)^2 \\ &= 0 \end{aligned}$$

Jadi, MSE $= var(\bar{X}) - (E(\bar{X}) - p)$
$= \frac{(1-p)}{np^2} - 0 = \frac{(1-p)}{np^2}$

### 3.4. Menentukan Distribusi Sampling Kolmogorov-smirnov dibawah $H_0$

*Tabel 1* ***Jumlah Lemparan Bola Basket 1***

| Orang Ke - | Jumlah lemparan |
|---|---|
| 1 | 14 |
| 2 | 12 |
| 3 | 11 |
| 4 | 10 |
| 5 | 7 |
| 6 | 9 |
| 7 | 5 |
| 8 | 8 |
| 9 | 6 |
| 10 | 7 |
| 11 | 9 |
| 12 | 10 |
| 13 | 6 |
| 14 | 11 |
| 15 | 7 |
| 16 | 5 |
| 17 | 9 |
| 18 | 12 |
| 19 | 8 |
| 20 | 9 |
| 21 | 5 |
| 22 | 7 |
| 23 | 8 |
| 24 | 10 |
| 25 | 8 |
| Rata-rata | 8,52 |

Data di atas kemudian disajikan dalam bentuk probabilitas untuk menggambarkan kemungkinan terjadinya suatu kejadian variabel $X$ dimana $\hat{p} = \frac{1}{8{,}52}$. Berikut adalah perhitungan $\hat{F}_n(x)$ untuk setiap kemungkinan jumlah lemparan persamaan untuk $\hat{F}_n(x) = \frac{i}{n}$

Dimana
$i$ = jumlah observasi kumulatif kurang dari
$n$ = banyaknya observasi
Dalam kasus pelemparan bola basket berikut perhitungan $\hat{F}_n(x)$

$\hat{F}_n(1) = 0$
$\hat{F}_n(2) = 0$
$\hat{F}_n(3) = 0$
$\hat{F}_n(4) = 0$
$\hat{F}_n(5) = \frac{3}{25} = 0{,}12$
$\hat{F}_n(6) = \frac{5}{25} = 0{,}08$
$\hat{F}_n(7) = \frac{9}{25} = 0{,}36$
$\hat{F}_n(8) = \frac{13}{25} = 0{,}52$
$\hat{F}_n(9) = \frac{17}{25} = 0{,}68$
$\hat{F}_n(10) = \frac{20}{25} = 0{,}8$
$\hat{F}_n(11) = \frac{22}{25} = 0{,}88$
$\hat{F}_n(12) = \frac{24}{25} = 0{,}96$
$\hat{F}_n(13) = 0$
$\hat{F}_n(14) = \frac{25}{25} = 1$

Setelah menghitung $\hat{F}_n(x)$ selanjutnya dihitung $F_0(x)$ dengan menerapkan rumus sebagai berikut

$$F_0(k) = 1 - (1-p)^k$$

dimana
$p =$ probabilitas keberhasilan dalam satu uji percobaan
$k$ = jumlah pelemparan
Dalam kasus pelemparan bola basket berikut perhitungan $F_0(x)$

$F_0(1) = 1-(1-p)^k = 1-\left(1-\frac{1}{8{,}52}\right)^1 = 0{,}117371$
$F_0(2) = 1-(1-p)^k = 1-\left(1-\frac{1}{8{,}52}\right)^2 = 0{,}220965$
$F_0(3) = 1-(1-p)^k = 1-\left(1-\frac{1}{8{,}52}\right)^3 = 0{,}312402$
$F_0(4) = 1-(1-p)^k = 1-\left(1-\frac{1}{8{,}52}\right)^4 = 0{,}393105$
$F_0(5) = 1-(1-p)^k = 1-\left(1-\frac{1}{8{,}52}\right)^5 = 0{,}464338$
$F_0(6) = 1-(1-p)^k = 1-\left(1-\frac{1}{8{,}52}\right)^6 = 0{,}527209$
$F_0(7) = 1-(1-p)^k = 1-\left(1-\frac{1}{8{,}52}\right)^7 = 0{,}582701$
$F_0(8) = 1-(1-p)^k = 1-\left(1-\frac{1}{8{,}52}\right)^8 = 0{,}631679$
$F_0(9) = 1-(1-p)^k = 1-\left(1-\frac{1}{8{,}52}\right)^9 = 0{,}67491$
$F_0(10) = 1-(1-p)^k = 1-\left(1-\frac{1}{8{,}52}\right)^{10} = 0{,}713066$
$F_0(11) =1-(1-p)^k = 1-\left(1-\frac{1}{8{,}52}\right)^{11} = 0{,}746743$
$F_0(12) = 1-(1-p)^k = 1-\left(1-\frac{1}{8{,}52}\right)^{12} = 0{,}776468$
$F_0(13) = 1-(1-p)^k = 1-\left(1-\frac{1}{8{,}52}\right)^{13} = 0{,}802704$
$F_0(14) = 1-(1-p)^k = 1-\left(1-\frac{1}{8{,}52}\right)^{14} = 0{,}825861$

Simulasi Monte Carlo adalah simulasi tipe probabilitas yang mendekati solusi sebuah masalah dari proses acak. Simulasi Monte Carlo melibatkan penetapan distribusi pengambilan sampel acak dari distribusi untuk menghasilkan data. Pada penelitian ini, yang menjadi fokus utama dari uji Kolmogorov-Smirnov adalah pengujian terhadap distribusi geometrik data, dimana hal ini terkait dengan kasus yang menjadi batasan dalam penelitian ini yaitu banyaknya lemparan yang masuk ke ring sampai sukses pertama pada permainan bola basket.

Distribusi sampling dari statistik $D_n$ dibawah kondisi $H_0$ dapat dihampiri dengan menerapkan simulasi Monte Carlo. Kuantil-kuantil dari distribusi $D_n$ dapat dihampiri dengan simulasi Monte Carlo untuk berbagai ukuran sampel $n$ dan berbagai tingkat signifikansi $\alpha$, dapat dihampiri dengan simulasi

Jurnal Matematika, Komputasi dan Statistika
Volume 4 Nomor 1, pp. 491-497, Januari-April 2024,
ISSN: 2503 - 2984

Monte Carlo menggunakan komputer. Hasil simulasi ditampilkan pada Tabel 2.

*Tabel 2* ***Tabel kuantil-kuantil dari statistika $D_n$***

| n | $\alpha$ | | | | |
|---|---|---|---|---|---|
| | 1% | 5% | 10% | ⋮ | 99% |
| 1 | 0,037 | 0,0478 | 0,0612 | ⋮ | 0,1965 |
| 2 | 0,0469 | 0,0805 | 0,1226 | ⋮ | 0,1752 |
| 3 | 0,1975 | 0,2092 | 0,2237 | ⋮ | 0,2477 |
| 4 | 0,372 | 0,3733 | 0,3748 | ⋮ | 0,4576 |
| 5 | 0,306 | 0,3147 | 0,3257 | ⋮ | 0,4024 |
| 6 | 0,3522 | 0,3947 | 0,4477 | ⋮ | 0,4944 |
| 7 | 0,2211 | 0,2313 | 0,244 | ⋮ | 0,3178 |
| 8 | 0,237 | 0,2535 | 0,2741 | ⋮ | 0,3205 |
| 9 | 0,4087 | 0,4111 | 0,4141 | ⋮ | 0,4873 |
| 10 | 0,3427 | 0,3769 | 0,4195 | ⋮ | 0,6443 |
| 11 | 0,2519 | 0,2955 | 0,3499 | ⋮ | 0,4719 |
| 12 | 0,3049 | 0,3248 | 0,3496 | ⋮ | 0,3897 |
| 13 | 0,3266 | 0,3392 | 0,3549 | ⋮ | 0,4621 |
| 14 | 0,3325 | 0,339 | 0,3472 | ⋮ | 0,3671 |
| 15 | 0,277 | 0,3103 | 0,352 | ⋮ | 0,4476 |
| 16 | 0,2788 | 0,2874 | 0,298 | ⋮ | 0,3385 |
| 17 | 0,2995 | 0,3331 | 0,3752 | ⋮ | 0,4 |
| 18 | 0,253 | 0,3043 | 0,3684 | ⋮ | 0,4457 |
| 19 | 0,4456 | 0,4515 | 0,459 | ⋮ | 0,474 |
| 20 | 0,3448 | 0,3805 | 0,4252 | ⋮ | 0,4801 |
| 21 | 0,3268 | 0,3552 | 0,3908 | ⋮ | 0,5324 |
| 22 | 0,4621 | 0,4754 | 0,492 | ⋮ | 0,6039 |
| 23 | 0,4582 | 0,4683 | 0,481 | ⋮ | 0,5093 |
| 24 | 0,3313 | 0,4019 | 0,4902 | ⋮ | 0,5281 |
| 25 | 0,3361 | 0,3667 | 0,405 | ⋮ | 0,4726 |
| 26 | 0,3819 | 0,3976 | 0,4173 | ⋮ | 0,4868 |
| 27 | 0,3434 | 0,3487 | 0,3552 | ⋮ | 0,5317 |
| 28 | 0,445 | 0,4743 | 0,511 | ⋮ | 0,5505 |
| 29 | 0,4587 | 0,4901 | 0,5295 | ⋮ | 0,6121 |
| 30 | 0,4574 | 0,4753 | 0,4976 | ⋮ | 0,6104 |

Kuantil-kuantil pendekatan dari $D_n$ untuk $\hat{p} = \frac{1}{8{,}52}$ dan beberapa $n$ diperoleh dengan menggunakan program, program tersebut ditulis dalam bahasa R. Pada simulasi ini sampel dibangkitkan dari distribusi geometrik dimana parameter $p$ diestimasi dari data pada tabel 1. Jadi sampel dibangkitkan dari distribusi geometrik $\left(\frac{1}{8{,}52}\right)$ dengan memilih 1000000 perulangan.

$$D_{13} = \sup_{x \in \mathbb{R}} \sqrt{n}|F_{13}(x) - F_0(x)|$$
$$= \sqrt{25}|0 - 8027|$$
$$= 4{,}0135$$

*Tabel 3* ***Nilai selisih*** $|F_n(x) - F_0(x)|$

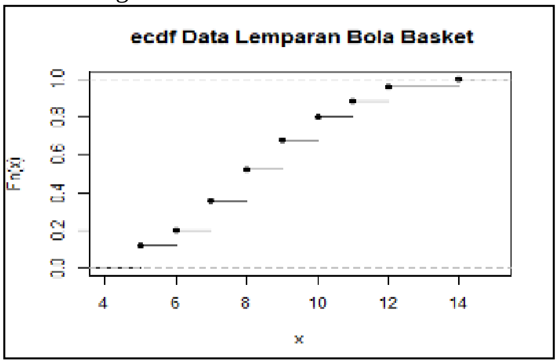
| n | $F_n(x)$ | $F_0(x)$ | D | \|D\| |
|---|---|---|---|---|
| 1 | 0 | 0,1174 | -0,5869 | 0,5869 |
| 2 | 0 | 0,2210 | -1,1048 | 1,1048 |
| 3 | 0 | 0,3124 | -1,5620 | 1,5620 |
| 4 | 0 | 0,3931 | -1,9655 | 1,9655 |
| 5 | 0,12 | 0,4643 | -1,7217 | 1,7217 |
| 6 | 0,2 | 0,5272 | -1,6360 | 1,6360 |
| 7 | 0,36 | 0,5827 | -1,1135 | 1,1135 |
| 8 | 0,52 | 0,6317 | -0,5584 | 0,5584 |
| 9 | 0,68 | 0,6749 | 0,0255 | 0,0255 |
| 10 | 0,8 | 0,7131 | 0,4347 | 0,4347 |
| 11 | 0,88 | 0,7467 | 0,6663 | 0,6663 |
| 12 | 0,96 | 0,7765 | 0,9177 | 0,9177 |
| 13 | 0 | 0,8027 | -4,0135 | 4,0135 |
| 14 | 1 | 0,8259 | 0,8707 | 0,8707 |

MSE mengukur seberapa dekat estimasi suatu parameter atau model dengan nilai yang sebenarnya. Dalam konteks distribusi geometrik, MSE mengevaluasi seberapa baik suatu estimasi parameter atau prediksi dalam distribusi ini.

$$\text{MSE} = \frac{(1-p)}{np^2} = \frac{\left(1-\frac{1}{8{,}52}\right)}{25\left(\frac{1}{8{,}52}\right)^2} = 2{,}5628$$

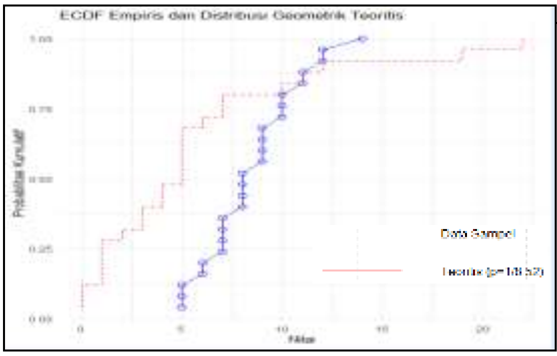
### 3.5. Analisis Data

Grafik ecdf dari data lemparan bola basket diatas adalah sebagai berikut :




*Gambar 1* ***Grafik ecdf dari data jumlah lemparan bola basket ke ring***

Untuk mengetahui secara visual maka dibuat grafik ECDF (*Empiris Cumulative Distribution Function*) dari distribusi geometrik dengan $p = \frac{1}{8{,}52}$ dan $n =$ satu sampai tak hingga diperoleh grafiknya sebagai berikut:




*Gambar 2* ***Grafik ECDF dari data sampel dan data distribusi Geometrik dengan*** $\boldsymbol{p = \frac{1}{8{,}52}}$

Dari grafik 2 pola grafik ECDF (*Empiris Cumulative Distribution Function*) dari data sampel mengikuti pola dari data populasi distribusi geometrik dengan $p = \frac{1}{8{,}52}$. Secara visual dapat disimpulkan bahwa data lemparan bola basket ke ring mengikuti pola distribusi geometrik.

Selanjutnya menentukan hipotesis mengenai distribusi peluang dari data yang akan ditentukan distribusinya. Dimana hipotesisnya adalah sebagai berikut:
$H_0$ : Populasi mengikuti distribusi Geometrik
$H_1$ : Populasi tidak mengikuti distribusi Geometrik

Berdasarkan tabel 2 nilai statistik uji kolmogorov-smirnov adalah $4{,}0135$ dengan nilai kuantil-kuantil $=\ 5\%$ , $n =\ 25$ adalah $0{,}3667$ hal ini menunjukan bahwa data jumlah lemparan bola basket ke ring mengikuti distribusi geometrik atau terima $H_0$.

## 4. Kesimpulan dan Saran

Berdasarkan hasil pembahasan yang dapat disimpulkan sebagai berikut:

Estimasi parameter $(P)$ distribusi geometrik dengan metode *MLE* diperoleh dengan persamaan $p = \frac{1}{\bar{x}}$. Ketakbiasan penduga *MLE* diuji dengan membandingakn nilai variansi dan nilai *BBCR*, diperoleh nilai var $\neq BBCR$ sehingga $\frac{1}{\bar{x}}$ bukan merupakn *UMVUE*.

Data lemparan bola basket mengikuti distribusi Geometrik dengan statistik uji Kolmogorov-Smirnov $(D)$ diperoleh sebesar $4{,}0135$. Untuk menginterpretasikan hasil ini, dibandingkan dengan nilai kritis dari tabel kuantil-kuantil. Tabel kuantil-kuantil diperoleh dari simulasi monte-carlo dengan $1000000$ perulangan dengan nilai $p = \frac{1}{8{,}52}$, untuk tingkat signifikansi $5\%$ ( $\alpha = 0{,}05$) dengan ukuran sampel $(n = 25)$. Nilai tabel kuantil-kuantil untuk $n = 25$ dan $\alpha = 0{,}05$ adalah $0{,}3667$. Karena nilai $D$ yang diperoleh $(4{,}0135)$ lebih besar dari nilai tabel kuantil-kuantil (0,3667), maka hipotesis nol yang menyatakan bahwa data mengikuti distribusi geometrik.

Berdasarkan kesimpulan di atas, maka penulis memberikan saran sebagai berikut:

Untuk penelitian selanjutnya, perlu mendalami dan mengkaji lebih jauh karakteristik masing-masing distribusi peluang dalam konteks kehidupan sehari-hari, dengan fokus pada aplikasi praktis dan potensi implementasi untuk pemahaman lebih mendalam terhadap fenomena yang diamati..

**Ucapan Terima Kasih**

Penelitian ini dapat terlaksana dengan baik atas bantuan dari berbagai pihak. Peneliti mengucapkan terima kasih yang sebesar-besarnya kepada pembimbing 1 Bapak Prof.Dr.rer.nat. Wayan Somayasa, S.Si., M.Si pembimbing 2, Alfian, S.Si., M.Sc., dan para penguji yang memberikan saran, kritikan dan ide sehingga penelitian dapat terlaksana dengan baik.

**Daftar Pustaka**

[1] Daniel, W. W. (1989). Statistika nonparametrik terapan (Terjemahan Alex Tri Kantjono W). *New York: Hounghyon Mifflin Company*.

[2] Rini, D. S. (2013). *Perbandingan Uji Goodness Of Fit untuk Normalitas antara Metode Klasik dan Bayesan*. etd.repository.ugm.ac.id. Retrieved from http://etd.repository.ugm.ac.id/home/detail_pencarian/62976

[3] Savapandit,R.R.,dan Gogoi, B. (2015). Bootstrap and Other Tests For Goodness of Fit.

[4] Hardina, D. R. (2016). GRAFIK PENGENDALI ATRIBUT BERBASIS DISTRIBUSI GEOMETRIK DENGAN ESTIMASI BATAS PENGENDALI (Doctoral dissertation, Universitas Gadjah Mada).

[5] Somayasa, W. (2023). *Pengantar Teori Peluang dan Statistika Matematika*. (Muhammad Kabil Djafar, Ed.). Yogyakarta: DEEPUBLISH.

[6] Fallo, J. O., Setiawan, A., dan Susanto, B. (2013). Uji Normalitas Berdasarkan Metode Anderson-Darling, Cramer-von Mises dan Liliefor Menggunakan Metode Bootstrap. *Prosiding*.

[7] Yanuar, F., Wulandari, S., & HG, I. R. (2021). Analisis Survival untuk Parameter Skala dari Distribusi Weibull Menggunakan Mle dan Metode Bayesian. BAREKENG: Jurnal Ilmu Matematika dan Terapan, 15(1), 147-156.