Epsilon: Jurnal Matematika Murni dan Terapan
http://ppjp.ulm.ac.id/journals/index.php/epsilon
p-ISSN: 1978-4422 | e-ISSN 2656-7660
Vol.18 No.2 (Desember 2024) Hal. 240-255

PENGGUNAAN RUMUS STIRLING DAN FUNGSI PEMBANGKIT MOMEN PADA PROSES PEMBUKTIAN DISTRIBUSI NORMAL SEBAGAI PENDEKATAN DISTRIBUSI BINOMIAL

Raden Gunawan Santosa[1], Nugroho Agus Haryono[2]

[1,2]Program Studi Informatika Fakultas Teknologi Informasi
Universitas Kristen Duta Wacana,
Jl. Dr. Wahidin Sudirohusodo 5-25 Yogyakarta 55224
email: *gunawan@staff.ukdw.ac.id*

ABSTRACT

Research in mathematics and statistics can be categorized into two main types: analytical-theoretical research and applied-practical research, both of which make significant contributions to statistical analysis. This research is particularly relevant to the analysis of large-scale binary variables, where the binomial distribution is often used as the basis for computation. In such cases, the normal distribution can serve as an approximation to the binomial distribution, reducing the complexity of calculations. This study aims to demonstrate that the normal distribution can be used as an approximation for the binomial distribution. Two proof methods are presented: one utilizing the moment-generating function and the other employing Stirling's formula. The methodology involves a literature review by gathering and analyzing various relevant references. The findings indicate that as the value of n n increases, the difference between calculations using the binomial and normal distributions approaches zero exponentially.

Keywords: Binomial distribution, normal distribution, distribution approximation, moment generating function, Stirling formula

ABSTRAK

Penelitian dalam bidang matematika dan statistika dapat dibagi menjadi dua jenis utama, yaitu penelitian analitis-teoritis dan penelitian aplikatif-praktis, yang keduanya memiliki kontribusi penting dalam analisis statistik. Penelitian ini memiliki relevansi khusus pada analisis variabel biner dalam skala besar, di mana distribusi binomial sering digunakan sebagai dasar komputasi. Dalam situasi semacam ini, distribusi normal dapat menggantikan distribusi binomial untuk mengurangi kompleksitas perhitungan. Penelitian ini bertujuan untuk membuktikan bahwa distribusi normal dapat digunakan sebagai pendekatan distribusi binomial. Dua proses pembuktian disajikan dalam penelitian ini, yaitu melalui penggunaan fungsi pembangkit momen dan rumus Stirling. Metode yang digunakan adalah studi literatur dengan mengumpulkan berbagai referensi yang relevan. Hasil penelitian menunjukkan bahwa semakin besar nilai n, perbedaan antara perhitungan menggunakan distribusi binomial dan distribusi normal akan mendekati nol secara eksponensial.

Kata kunci: Distribusi binomial, distribusi normal, pendekatan distribusi, fungsi pembangkit momen, rumus stirling

Received: 22 November 2024, Accepted: 01 Januari 2025, Published: 3 Januari 2025

PENDAHULUAN

Distribusi binomial merupakan salah satu distribusi variabel acak diskret yang sering digunakan dalam aplikasi pengolahan data statistik. Distribusi ini, bersama dengan distribusi normal, memiliki peran penting dalam berbagai aspek pengolahan data statistik, terutama dalam analisis probabilitas, inferensi statistik, dan pengambilan keputusan.

Distribusi binomial digunakan untuk memodelkan fenomena yang melibatkan dua hasil yang saling eksklusif dalam satu percobaan, seperti sukses atau gagal, ya atau tidak. Distribusi ini memiliki sejumlah keunggulan yang menjadikannya penting dalam analisis statistik. Pertama, distribusi binomial merupakan model probabilitas diskret yang berfungsi sebagai alat utama untuk menganalisis data berbentuk diskret, seperti menghitung jumlah keberhasilan dalam n percobaan. Kedua, distribusi ini memiliki aplikasi yang luas dan dapat diterapkan dalam berbagai bidang. Contohnya, dalam riset pemasaran, distribusi binomial dapat digunakan untuk menganalisis jumlah pelanggan yang membeli suatu produk. Dalam uji hipotesis, distribusi ini digunakan untuk menguji proporsi suatu populasi, sedangkan dalam genetika, distribusi ini membantu dalam analisis distribusi alel tertentu. Ketiga, distribusi binomial menjadi dasar penting dalam inferensi statistik. Hasil distribusi ini sering dimanfaatkan untuk mengestimasi parameter populasi, seperti proporsi keberhasilan (p), yang menjadi salah satu aspek fundamental dalam pengambilan keputusan berbasis data. Distribusi Binomial adalah distribusi variabel acak diskret yang sering digunakan pada aplikasi pengolahan data statistik. Distribusi binomial dan distribusi normal memiliki peran penting dalam pengolahan data statistik, khususnya dalam analisis probabilitas, inferensi statistik, dan pengambilan keputusan.

Distribusi normal, yang juga dikenal sebagai distribusi Gauss, merupakan distribusi probabilitas kontinu dengan peran penting dalam statistik dan analisis data. Kepentingannya didasarkan pada beberapa sifat utama, yaitu kemunculannya secara alami dalam berbagai fenomena alam, sosial, dan ekonomi, seperti tinggi badan, skor ujian, atau fluktuasi pasar. Teorema Limit Tengah (Central Limit Theorem) menyatakan bahwa data dari berbagai distribusi cenderung mengikuti distribusi normal jika ukuran sampel cukup besar, menjadikannya dasar dalam banyak analisis statistik. Aplikasinya meliputi inferensi statistik, seperti pembuatan interval kepercayaan, pengujian hipotesis, dan analisis regresi. Sifat simetri dan keberadaan fungsi standar pada distribusi normal juga mempermudah penghitungan probabilitas serta analisis data, sehingga menjadi alat yang efisien dalam berbagai bidang penelitian dan praktik statistik.

Distribusi binomial dapat didekati dengan distribusi normal jika jumlah percobaan (n) besar dan probabilitas sukses (p) tidak terlalu kecil atau besar, sesuai dengan aturan $np \geq 0{,}5$ dan $n(1-p) \geq 0{,}5$. Hal ini memungkinkan penggunaan distribusi normal sebagai pendekatan untuk distribusi binomial dalam kasus tertentu, yang sering digunakan untuk mempermudah analisis. Keduanya adalah alat penting dalam analisis data karena kemampuan mereka untuk memodelkan fenomena dunia nyata dan mendukung pengambilan keputusan berbasis data. Distribusi binomial cocok untuk data diskrit dengan dua kemungkinan hasil, sementara distribusi normal lebih relevan untuk data kontinu yang tersebar secara simetris.

Berbagai penelitian telah memanfaatkan aplikasi distribusi binomial untuk menganalisis berbagai kasus dalam kehidupan nyata. Kebijakan satu anak yang diterapkan di China sejak era Deng Xiaoping pada tahun 1979 hingga akhirnya dihapuskan pada akhir 2015 menjadi salah satu objek penelitian Hermawan et al., (2019). Mereka menemukan bahwa jika peluang kelahiran bayi laki-laki tinggi, kebijakan tersebut dapat berhasil, namun akhirnya digantikan oleh kebijakan yang memperbolehkan setiap pasangan memiliki dua anak. Setyaningsih et al., (2021) mengkaji probabilitas keberhasilan pemerintah dalam mengatasi kemacetan di Jakarta sebagai Ibu Kota melalui kuesioner kepada teman dan kerabat mahasiswa Teknik Industri Universitas Indraprasta PGRI. Selanjutnya, Pebriyanto et al., (2021) menggunakan distribusi binomial untuk mengukur tingkat kegagalan produksi tas di home industri @One homemade, dengan hasil bahwa tingkat kegagalan produksi sangat kecil. Terakhir, Loban et al., (2023) memanfaatkan model distribusi binomial untuk mengestimasi probabilitas kesuksesan masa studi mahasiswa di Program Studi Matematika Universitas Tribuana Kalabahi. Dengan data 105 lulusan periode 2017-2021, mereka menemukan bahwa rata-rata kesuksesan mahasiswa menyelesaikan kuliah dalam waktu ≤8 semester tidak sama dengan 0,50. Penelitian-penelitian ini menunjukkan fleksibilitas distribusi binomial dalam berbagai konteks analisis probabilitas.

Tujuan mengimplementasikan distribusi binomial adalah untuk mengestimasi probabilitas sukses pada uji coba kualitas layanan sistem informasi dengan menggunakan variabel-variabel yang ada pada model E-ServQual. Penemuan penelitian ini bahwa jumlah sampel kecil dan nilai probabilitas yang besar akan meningkatkan probabilitas kesuksesan yang akan diperoleh (Diana, 2017). Sedangkan beberapa penelitian teoritis yang terkait dengan distribusi Binomial, diantaranya, (Sitopu & Siswadi, 2022), berfokus pada perumusan model distribusi probabilitas dan *Fungsi Pembangkit Momen* (*FPM*) variabel acak diskrit. Berdasarkan hasil dari kajian tersebut, diperoleh model distribusi probabilitas dan fungsi pembangkit momen variabel acak Distribusi Binomial, Distribusi Binomial

Negatif, Distribusi Geometrik, Distribusi Poisson, Distribusi Trinomial dan Distribusi Multinomial. Penelitin yang dilakukan oleh Sitopu hanya menyajikan bentuk-bentuk FPM dari bermacam-macam variabel acak. Berdasarkan ide penelitian ini, maka peneliti tertarik untuk melakukan dan menjabarkan kegunaan dari FPM.

Kurniawati dan Musadi (2023) membuktikan pendekatan distribusi normal melalui distribusi binomial dan distribusi Poisson baik secara teoritis maupun induktif. Secara teoritis, pendekatan tersebut dibuktikan menggunakan metode pembangkit momen, sedangkan secara induktif, pendekatan ini divisualisasikan melalui grafik distribusi binomial dan distribusi Poisson, yang menunjukkan kecenderungan mendekati kurva normal ketika jumlah data sangat besar atau mendekati tak hingga. Berbeda dengan penelitian Musadi dan Kurniawati, penelitian ini memiliki relevansi dalam menyoroti proses pembuktian secara lugas dengan memanfaatkan dua pendekatan yang berbeda.

Artikel penelitian ini berfokus pada distribusi Binomial dan distribusi Normal. Variabel acak X, yang mewakili jumlah keberhasilan dalam n percobaan pada eksperimen binomial, disebut sebagai variabel acak binomial, sedangkan distribusi probabilitas X dalam eksperimen tersebut dikenal sebagai distribusi Binomial. Dalam rumus Binomial, n merepresentasikan jumlah total percobaan, sedangkan X menunjukkan jumlah total keberhasilan. Selisih antara jumlah total percobaan dan jumlah total keberhasilan, yaitu $n - x$, mencerminkan jumlah total kegagalan dalam n percobaan. Kedua distribusi ini memiliki bentuk fungsional yang dinyatakan dalam persamaan (1) dan (2) (Walck, 2007).

Bentuk distribusi Binomial :

$$P(X = x) = \binom{n}{x} p^x q^{n-x} \tag{1}$$
$$x = 0, 1, 2, 3, \dots, n$$
$$q = 1 - p$$

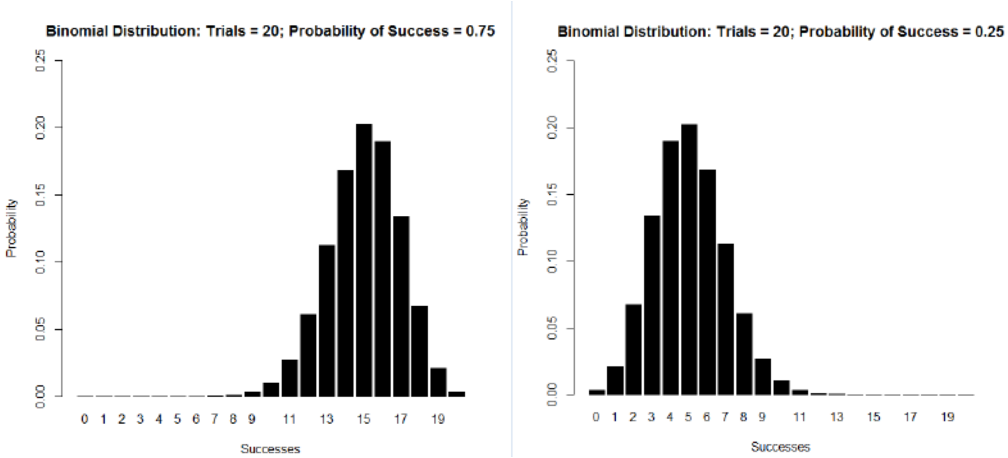
Distribusi Binomial ini juga mempunyai mean = np dan variansi =npq.

Distribusi probabilitas Normal, jika digambarkan secara visual, akan berbentuk lonceng dengan beberapa sifat, yaitu total luasan di bawah kurva adalah 1, kurvanya simetris terhadap nilai mean (μ), ekor kurva bagian kiri dan kanan akan memanjang tanpa batas. Bentuk distribusi normal dengan mean μ dan standar deviasi σ:

$$f(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] \quad ; -\infty < x < \infty \tag{2}$$

Distribusi Binomial mempunyai dua parameter yaitu n dan p , sedangkan distribusi normal mempunyai parameter μ dan σ.

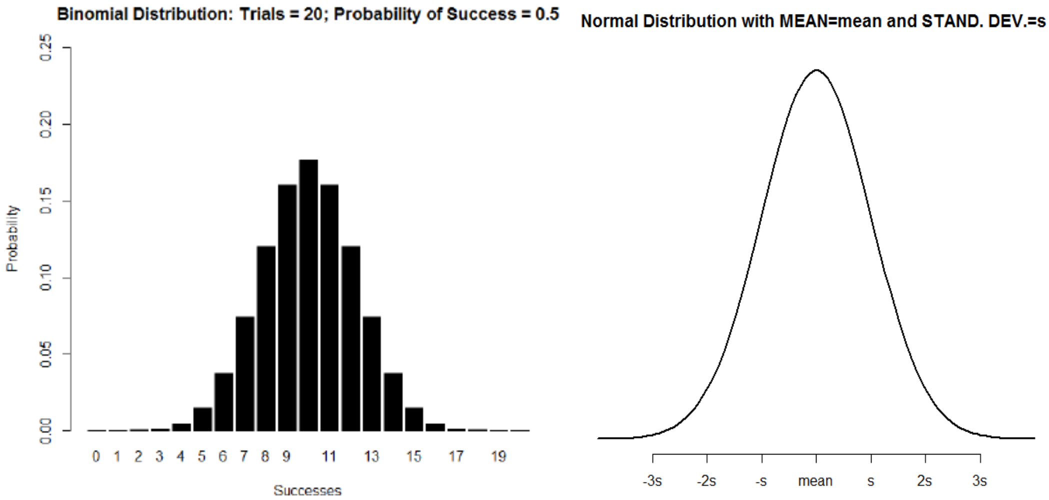
Pada Gambar 1(a) disajikan grafik distribusi binomial dengan $n = 20$ dan $p = 0.75$, sedangkan Gambar 1(b) merupakan gambar visual distribusi binomial dengan $n = 20$ dan $p = 0.25$.



(a) (b)

Gambar 1. (a) Binomial dengan n=20 dan p=0.75 (b) Binomial dengan n=20 dan p=0.25

Terlihat dalam Gambar 1(a) dan dan 1(b), bahwa untuk besar n yang sama dan $p > 0.50$ maka distribusi binomial akan menceng kekiri (ekor di kiri), sedangkan untuk $p < 0.50$ maka distribusi ini akan menceng ke kanan (ekor di kanan).



Gambar 2. (a) Binomial dengan n=20 dan p=0.50 (b) Normal dengan μ=mean dan σ=s

Pada Gambar 2(a) adalah gambar distribusi Binomial dengan $n = 20$ dan $p = 0.50$, sedangkan Gambar 2(b) merupakan gambar visual distribusi Normal dengan $\mu = mean$ dan $\sigma = s$. Distribusi Binomial dengan $n = 20$ dan $p = 0.50$, berbentuk simetris. Secara teori, distribusi Binomial dibangun dari variabel acak diskrit, sedangkan distribusi Normal adalah termasuk variabel acak yang berbentuk kontinu.

Menurut Mann (2020), secara praktis distribusi normal dapat digunakan sebagai pendekatan terhadap distribusi binomial jika memenuhi syarat $np > 5$ dan

$nq > 5$. Karena perbedaan kategori antara kedua distribusi tersebut, di mana distribusi normal digunakan untuk variabel acak kontinu sedangkan distribusi binomial untuk variabel acak diskrit. Pendekatan ini memerlukan faktor koreksi kontinuitas. Faktor koreksi kontinuitas melibatkan penambahan 0,5 pada batas atas nilai x dan pengurangan 0,5 pada batas bawah nilai x ketika distribusi normal digunakan untuk memperkirakan distribusi binomial.

Tujuan dari penelitian ini adalah untuk menjabarkan proses pembuktian distribusi normal sebagai pendekatan distribusi binomial dengan menggunakan FPM dan Rumus Stirling.

TINJAUAN PUSTAKA

Fungsi Pembangkit Momen (Moment Generating Function) adalah fungsi yang penting dalam pembentukan momen-momen suatu distribusi (Walpole et al., 2012). Karena ada dua kelas variabel acak, yaitu variabel acak diskrit dan variabel acak kontinu, maka Fungsi Pembangkit Momen juga harus mengikuti kedua kelas variabel acak tersebut. Untuk variabel acak X yang mempunyai distribusi probabilitas diskrit $P(X = x)$, maka Fungsi Pembangkit Momen dari X adalah

$$M_x(t) = E(e^{tx}) = \sum e^{tx} P(X = x) \tag{3}$$

Sedangkan, untuk variabel acak X yang mempunyai distribusi probabilitas kontinu f(x), maka Fungsi Pembangkit Momen dari X adalah

$$M_x(t) = E(e^{tx}) = \int_{-\infty}^{\infty} e^{tx} f(x)dx \tag{4}$$

Sebelum proses pembuktian dengan Fungsi Pembangkit Momen dan rumus Stirling, maka diberikan terlebih dahulu lima teorema yang sangat berkaitan dengan kedua proses pembuktian tersebut.

Teorema 1. Fungsi Pembangkit Momen dari distribusi Binomial berbentuk:

$$M_x(t) = (\, q + pe^t)^n \tag{5}$$

Bukti.

$$M_x(t) = E(e^{tx}) = \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x q^{n-x} = \sum_{x=0}^{n} \binom{n}{x} e^{tx} p^x q^{n-x} =$$
$$\sum_{x=0}^{n} \binom{n}{x} p^x e^{tx} q^{n-x} = \sum_{x=0}^{n} \binom{n}{x} (pe^t)^x q^{n-x} = (\, q + pe^t)^n$$

Teorema 2. Fungsi Pembangkit Momen dari distribusi normal, berbentuk:

$$M_x(t) = \; e^{\mu t + \frac{(\sigma^2 t^2)}{2}} \tag{6}$$

Bukti.

$$M_x(t) = \int_{-\infty}^{\infty} e^{tx} f(x)dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tx} exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] dx$$

Misalkan $z = \frac{X-\mu}{\sigma}$ maka $x = \mu + \sigma z$ dan $dz = dx/\sigma$, sehingga

$$M_x(t) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tx} exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{t(\mu + \sigma z)} exp\left[-\frac{1}{2}(z)^2\right] dz$$

$$= \frac{1}{\sqrt{2\pi}}\, e^{\mu t}\int_{-\infty}^{\infty} e^{\sigma z t} exp\left[-\frac{1}{2}(z)^2\right] dz$$
$$= \frac{1}{\sqrt{2\pi}}\, e^{\mu t}\int_{-\infty}^{\infty} exp\left[-\frac{1}{2}(z^2 - 2\sigma z t)\right] dz$$
$$= \frac{1}{\sqrt{2\pi}}\, e^{\mu t}\int_{-\infty}^{\infty} exp\left[-\frac{1}{2}(z^2 - 2\sigma t z + \sigma^2 t^2 - \sigma^2 t^2)\right] dz$$
$$= \frac{1}{\sqrt{2\pi}}\, e^{\mu t}\int_{-\infty}^{\infty} exp\left[-\frac{1}{2}(z^2 - 2\sigma t z + \sigma^2 t^2) + \frac{(\sigma^2 t^2)}{2}\right] dz$$
$$= \frac{1}{\sqrt{2\pi}}\, e^{\mu t}\exp\left[\frac{(\sigma^2 t^2)}{2}\right]\int_{-\infty}^{\infty} exp\left[-\frac{1}{2}(z^2 - 2\sigma t z + \sigma^2 t^2)\right] dz$$
$$= \frac{1}{\sqrt{2\pi}}\, e^{\mu t + \frac{(\sigma^2 t^2)}{2}}\int_{-\infty}^{\infty} exp\left[-\frac{1}{2}(z - \sigma t)^2\right] dz$$

Misalkan $u = z - \sigma t$ maka $du = dz$

$$M_x(t) = \frac{1}{\sqrt{2\pi}}\, e^{\mu t + \frac{(\sigma^2 t^2)}{2}}\int_{-\infty}^{\infty} exp\left[-\frac{1}{2}(u)^2\right] du$$
$$= \frac{1}{\sqrt{2\pi}}\, e^{\mu t + \frac{(\sigma^2 t^2)}{2}}\sqrt{2\pi}$$
$$M_x(t) = e^{\mu t + \frac{(\sigma^2 t^2)}{2}}$$

Jika $X \sim N(\mu, \sigma^2)$ maka $Z = \frac{X-\mu}{\sigma}$ adalah distribusi Normal dengan mean 0 dan standar deviasi 1 ditulis sebagai Z ~ N(0,1). Sehingga bentuk distribusi $N(0,1)$ adalah $\emptyset(z) = \frac{1}{\sqrt{2\pi}} exp\,(-z^2/2)$ untuk $-\infty < z < \infty$ sebagai fungsi padat peluang distribusi normal baku (Hogg et al., 2018).

Teorema 3. Fungsi Pembangkit Momen distribusi normal baku, berbentuk:

$$M_z(t) = e^{\frac{t^2}{2}} \tag{7}$$

Bukti. nilai $\mu = 0$ dan $\sigma = 1$ disubstitusi ke Persamaan (6) diperoleh

$$M_z(t) = e^{0t + \frac{(1^2 t^2)}{2}} = e^{\frac{t^2}{2}}$$

Teorema 4. Jika $M_x(t)$ adalah Fungsi Pembangkit Momen variabel acak X dan $Z = \frac{X-\mu}{\sigma}$, maka *MGF* variabel acak Z adalah $M_z(t) = e^{-\mu t/\sigma} M_x(t/\sigma)$ (8)

Bukti. $M_z(t) = E(e^{tz}) = E\left(e^{t\left(\frac{X-\mu}{\sigma}\right)}\right) = E\left(e^{\frac{Xt-\mu t}{\sigma}}\right) = E\left(e^{\frac{Xt-\mu t}{\sigma}}\right) =$
$E\left(e^{\frac{Xt}{\sigma}} e^{\frac{-\mu t}{\sigma}}\right) = e^{-\mu t/\sigma} M_x(t/\sigma)$

Jadi $M_z(t) = e^{-\mu t/\sigma} M_x(t/\sigma)$

Teorema 5. $n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$ untuk $n \to \infty$ (9)

Bukti. Rumus Stirling didasarkan pada fungsi Gamma:

$$\Gamma(n+1) = n! = \int_0^{\infty} x^n e^{-x} dx$$
$$n! = \int_0^{\infty} e^{\ln(x^n) - x} dx = \int_0^{\infty} e^{\mathrm{n}\ln(x) - x} dx$$

Misalkan $x = n + y$, maka menjadi $n! = \int_0^\infty e^{\mathrm{n}\ln(n+y)-n-y}dy =$ $e^{-n}\int_0^\infty e^{\mathrm{n}\ln(n+y)-y}dy = e^{-n}\int_0^\infty e^{\mathrm{n}\ln \mathrm{n}+\mathrm{n}\ln(1+\frac{\mathrm{y}}{\mathrm{n}})-y}dy$. Karena $(n+y) = n(1+y/n)$, maka $n! = \left(\frac{n}{e}\right)^n \int_0^\infty e^{\mathrm{n}\ln(1+\frac{\mathrm{y}}{\mathrm{n}})-y}dy$.

Diketahui bahwa:

$\ln(1+\mathrm{u}) = u - \frac{u^2}{2} + \frac{u^3}{3} - \frac{u^4}{4} + \cdots$ dan $\ln\left(1+\frac{\mathrm{y}}{\mathrm{n}}\right) = \frac{y}{n} - \frac{y^2}{2n^2} + \frac{y^3}{3n^3} - \frac{y^4}{4n^4} + \cdots$, maka

$$n! = \left(\frac{n}{e}\right)^n \int_0^\infty e^{\mathrm{n}\ln(1+\frac{\mathrm{y}}{\mathrm{n}})-y}dy$$

$$n! = \left(\frac{n}{e}\right)^n \int_0^\infty e^{n\left(\frac{y}{n}-\frac{y^2}{2n^2}+\frac{y^3}{3n^3}-\frac{y^4}{4n^4}\cdots\right)-y}dy = \left(\frac{n}{e}\right)^n \int_0^\infty e^{\left(y-\frac{y^2}{2n}+\frac{y^3}{3n^2}-\frac{y^4}{4n^3}\cdots\right)-y}dy$$

$n! = \left(\frac{n}{e}\right)^n \int_0^\infty e^{\left(-\frac{y^2}{2n}+\frac{y^3}{3n^2}-\frac{y^4}{4n^3}\cdots\right)}dy$, selanjutnya, misalkan $y = v\sqrt{n}$, maka didapat $n! = \left(\frac{n}{e}\right)^n \sqrt{n}\int_{-\sqrt{n}}^\infty e^{\left(-\frac{v^2}{2}+\frac{v^3}{3\sqrt{\mathrm{n}}}-\cdots\right)}dv$. Untuk $n \to \infty$, maka bentuk yang memuat penyebut n dapat diabaikan, sehingga $n! = \left(\frac{n}{e}\right)^n \sqrt{n}\int_{-\sqrt{n}}^\infty e^{\left(-\frac{v^2}{2}+\theta(\sqrt{\mathrm{n}})\right)}dy$. Untuk $n \to \infty$ maka $\theta(\sqrt{\mathrm{n}}) \to 0$, $n! = \left(\frac{n}{e}\right)^n \sqrt{n}\int_{-\sqrt{n}}^\infty e^{\left(-\frac{v^2}{2}+\theta(\sqrt{\mathrm{n}})\right)}dy \approx$ $\left(\frac{n}{e}\right)^n \sqrt{n}\int_{-\infty}^\infty e^{\left(-\frac{v^2}{2}\right)}dy \approx \left(\frac{n}{e}\right)^n \sqrt{n}\sqrt{2\pi}$. Jadi terbukti bahwa $n! \approx$ $\left(\frac{n}{e}\right)^n \sqrt{n}\sqrt{2\pi} = \left(\frac{n}{e}\right)^n \sqrt{2\pi n}$.

METODE PENELITIAN

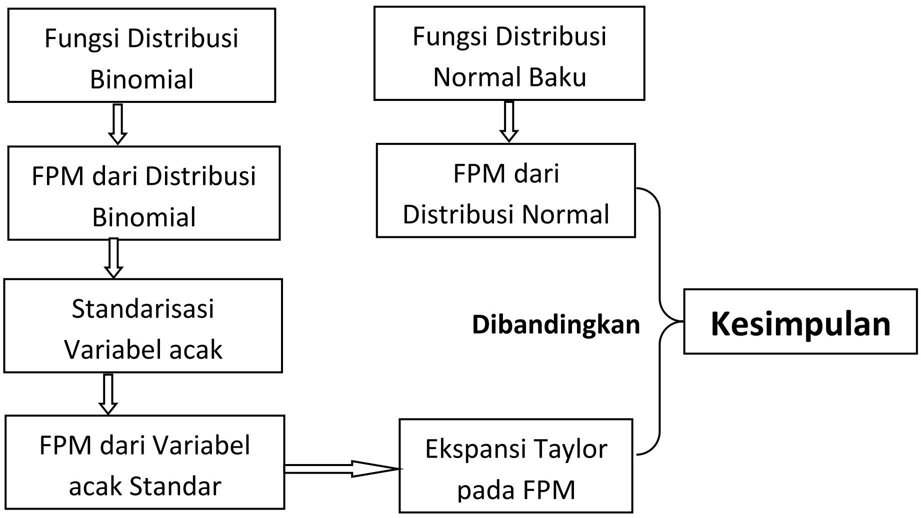
Metode penelitian yang digunakan pada penelitian ini adalah studi literatur yang terkait dengan distribusi binomial dan distribusi normal. Untuk proses pembuktian pertama digunakan Fungsi Pembangkit Momen (*Moment Generating Function*). Secara garis besar akan dibuktikan bahwa Moment Generating Function dari distribusi binomial akan mendekati distribusi normal baku untuk $n\to\infty$. Sedangkan cara pembuktian kedua digunakan rumus Stirling. Rumus Stirling ini merupakan rumus pendekatan untuk n faktorial $(n!)$, bila $n \to\infty$.

Ringkasan langkah-langkah proses pembuktian pendekatan distribusi normal untuk distribusi binomial dengan FPM disajikan sebagai berikut :

(a). Nyatakan bentuk fungsi distribusi binomial
(b). Definisikan FPM dari distribusi Binomial
(c). Standarisasi variabel acak distribusi binomial
(d). Nyatakan FPM variabel acak distribusi binomial yang sudah distandarisasi
(e). Lakukan ekspansi Taylor pada FPM distribusi variabel acak distribusi binomial yang sudah distandarisasi tersebut
(f). Bandingkan dengan FPM distribusi normal baku.

(g). Kesimpulan

Ilustrasi langkah-langkah pembuktian menggunakan FPM dapat dilihat dengan lebih jelas pada Gambar 3.

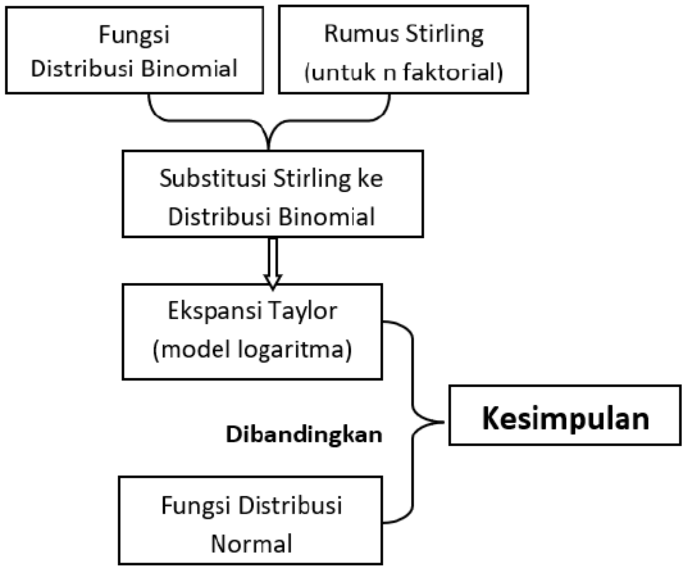


Gambar 3. Diagram Alir Proses Pembuktian dengan FPM

Ringkasan langkah-langkah proses pembuktian pendekatan distribusi normal untuk distribusi binomial, dengan rumus *Stirling* ,disajikan sebagai berikut :

(1) Nyatakan bentuk fungsi distribusi binomial.
(2) Nyatakan rumus *Stirling* untuk pendekatan n faktorial.
(3) Gabungkan (substitusi) ke dalam fungsi distribusi binomial.
(4) Sederhanakan dengan ekspansi Taylor untuk logaritma.
(5) Bandingkan dengan distribusi normal
(6) Kesimpulan

Ilustrasi langkah-langkah pembuktian menggunakan rumus *Stirling* dapat dilihat dengan lebih jelas pada Gambar 4.



Gambar 4. Diagram Alir Proses Pembuktian dengan rumus *Stirling*

HASIL DAN PEMBAHASAN

Pada bagian ini akan diberikan dua buah proses penjabaran secara matematis.

1. Pembuktian menggunakan Fungsi Pembangkit Momen (FPM)

Pada bagian ini akan diberikan cara pembuktian yang pertama, yaitu dengan menggunakan Fungsi Pembangkit Momen (FPM). Dari variabel acak binomial baku $Z = \frac{X-np}{\sqrt{npq}}$, akan ditunjukkan bahwa Fungsi Pembangkit Momennya akan menuju $exp(t^2/2)$ untuk $n \rightarrow \infty$.
Diketahui: jika $X \sim B(n,p)$, maka $M_x(t) = (\, q + pe^t)^n$,
Diberikan $\mu = np$ dan $\sigma^2 = npq$ (dengan $q = 1 - p$) adalah mean dan variansi distribusi binomial, maka dari persamaan (8):

$$\begin{aligned} M_z(t) &= e^{-\mu t/\sigma} M_x(t/\sigma) \\ &= e^{-npt/\sqrt{npq}}(\, q + pe^{t/\sqrt{npq}})^n \\ &= \left[e^{-pt/\sqrt{npq}}(\, q + pe^{t/\sqrt{npq}})\right]^n \\ &= \left[qe^{-pt/\sqrt{npq}} + pe^{qt/\sqrt{npq}})\right]^n \end{aligned}$$

Menggunakan ekspansi Taylor untuk e^x, yaitu:

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots \quad \text{maka}$$

$$M_z(t) = \left[q\left\{1 - \frac{pt}{\sqrt{npq}} + \frac{p^2t^2}{2npq} + \theta'(n^{-\frac{3}{2}})\right\} + p\left\{1 + \frac{qt}{\sqrt{npq}} + \frac{q^2t^2}{2npq} + \theta''(n^{-\frac{3}{2}})\right\})\right]^n$$

dengan $\theta'\left(n^{-\frac{3}{2}}\right)$ dan $\theta''(n^{-\frac{3}{2}})$ memuat bentuk $n^{-3/2}$ sebagai pangkat tertinggi dalam penyebutnya.

$$M_z(t) = \left[(p + q) + \frac{t^2 pq}{2npq}(p+q) + \theta(n^{-\frac{3}{2}})\right]^n$$
$$= \left[1 + \frac{t^2}{2n} + \theta(n^{-\frac{3}{2}})\right]^n$$

dimana $\theta(n^{-\frac{3}{2}})$ memuat bentuk $n^{-3/2}$ sebagai pangkat tertinggi dalam penyebutnya.

$$log\ M_z(t) = log\ \left[1 + \frac{t^2}{2n} + \theta(n^{-\frac{3}{2}})\right]^n$$

$$= n\ log\ \left[1 + \frac{t^2}{2n} + \theta(n^{-\frac{3}{2}})\right]$$
$$= n\ \left[\left\{\frac{t^2}{2n} + \theta(n^{-\frac{3}{2}})\right\} - \frac{1}{2}\left\{\frac{t^2}{2n} + \theta(n^{-\frac{3}{2}})\right\}^2 + \ ...\right]$$
$$log\ M_z(t) = \frac{t^2}{2} + \theta'''(n^{-\frac{1}{2}})$$

dimana $\theta(n^{-\frac{1}{2}})$ memuat bentuk $n^{-1/2}$ sebagai pangkat tertinggi dalam penyebutnya. Dengan n $\rightarrow\infty$ maka diperoleh:

$$\lim_{n\rightarrow\infty}\ log\ M_z(t) = \frac{t^2}{2}$$
$$\lim_{n\rightarrow\infty}\ M_z(t) = exp\ (\frac{t^2}{2})$$

Dapat ditarik kesimpulan bahwa bentuk di atas adalah merupakan Fungsi Pembangkit Momen dari distribusi normal baku. Dengan melihat kesamaan atau keunikan *moment generating function* maka variabel acak binomial standar akan menuju distribusi normal standar untuk $n\rightarrow\infty$. Dengan kata lain distribusi binomial menuju distribusi normal untuk $n \rightarrow\infty$.

2. Pembuktian menggunakan Rumus Stirling

Distribusi Normal sebagai bentuk pendekatan yang lain dari distribusi Binomial dibawah syarat-syarat sebagai berikut:

(a). Parameter n sebagai banyaknya trial z (percobaan dari distribusi Binomial menuju tak terhingga), yaitu $n\rightarrow\infty$.

(b). Nilai p dan q sangat kecil.

Distribusi Binomial dengan parameter n dan p diberikan dengan bentuk:

$P(X = x) = \binom{n}{x} p^x q^{n-x} = \frac{n!}{x!(n-x)!} p^x q^{n-x}$, x = 0, 1, 2, 3, …, n dan q = 1- p (10)

Misalkan kita punya bentuk binomial baku sebagai berikut:

$$Z = \frac{X-E(x)}{\sqrt{Var(x)}} = \frac{x-np}{\sqrt{npq}} \quad \text{untuk } x = 0,1,2,3,\dots,n \text{ dan } q = 1- p$$

maka:

bila $x = 0$ maka $z = \frac{x-np}{\sqrt{npq}} = \frac{0-np}{\sqrt{npq}} = -\sqrt{\frac{np}{q}}$ dan

bila $x = n$ maka $z = \frac{x-np}{\sqrt{npq}} = \frac{n-np}{\sqrt{npq}} = \sqrt{\frac{nq}{p}}$ (11)

Sehingga limit untuk n $\rightarrow\infty$, z akan mengambil nilai -∞ sampai dengan ∞ . Maka distribusi x akan merupakan distribusi kontinu dengan $-\infty < x < \infty$.

Dengan dua syarat di atas kita gunakan pendekatan rumus *Stirling* untuk $r!$ yaitu

$$\lim_{r\to\infty} r! = \sqrt{2\pi}\, e^{-r} r^{r+(\frac{1}{2})}$$

Kita mempunyai limit $n \to \infty$, sehingga akibatnya $x \to \infty$. Karena distribusi probabilitas binomial memuat variabel n dan x, maka nilai $lim\, p(x)$ diberikan sebagai berikut.

$$\lim p(x) = \lim \left[\frac{\sqrt{2\pi}\, e^{-n} n^{n+(\frac{1}{2})} p^x q^{n-x}}{\sqrt{2\pi}\, e^{-x} x^{x+(\frac{1}{2})} \sqrt{2\pi}\, e^{-(n-x)} r(n-x)^{(n-x)+(\frac{1}{2})}.}\right]$$

$$\lim p(x) = \lim \left[\frac{1}{\sqrt{2\pi}\sqrt{npq}} \frac{(np)^{x+1/2}(nq)^{n-x+1/2}}{x^{x+(\frac{1}{2})}\ (n-x)^{(n-x)+(\frac{1}{2})}.}\right]$$

$$\lim p(x) = \lim \left[\frac{1}{\sqrt{2\pi}\sqrt{npq}} \left(\frac{np}{x}\right)^{x+1/2} \left(\frac{nq}{n-x}\right)^{n-x+1/2}\right]$$

Dari persamaan (11), diperoleh $x = np + z\sqrt{npq}$ maka $\frac{x}{np} = 1 + z\sqrt{\frac{q}{np}}$.

Kemudian diperoleh $n - x = n - np - z\sqrt{npq} = nq - z\sqrt{npq}$

Jadi $\frac{n-x}{nq} = 1 - z\sqrt{\frac{p}{nq}}$ juga $\frac{dz}{dx} = \frac{1}{\sqrt{npq}}$ dan $dz = \frac{1}{\sqrt{npq}} dx$ (12)

Derivatif z dari Persamaan (12) diberikan oleh

dG(z) = g(z)dz = $\lim_{n\to\infty} \left[\frac{1}{\sqrt{2\pi}} \frac{1}{N}\right] dz$, dengan $N = \left(\frac{x}{np}\right)^{x+1/2} \left(\frac{n-x}{nq}\right)^{n-x+1/2}$ dan

$$logN = (x + 1/2) log\left(\frac{x}{np}\right) + (n - x + 1/2) log\left(\frac{n-x}{nq}\right)$$

$$= \left(np + z\sqrt{npq} + \frac{1}{2}\right) log\left(1 + z\sqrt{\frac{q}{np}}\right)$$

$$+\left(nq - z\sqrt{npq} + 1/2\right) log\left(1 - z\sqrt{\frac{p}{nq}}\right)$$

Dengan menggunakan ekspansi Taylor untuk log, yaitu

$$\log(1 + y) = y - \frac{y^2}{2} + \frac{y^3}{3} - \frac{y^4}{4} + \ ...$$

$$\log N = \left(np + z\sqrt{npq} + 1/2\right)\left[z\sqrt{\frac{q}{np}} - \frac{1}{2}z^2\left(\frac{q}{np}\right) + \frac{1}{3}z^3\left(\frac{q}{np}\right)^{3/2} - \cdots\right]$$

$$+ \left(nq - z\sqrt{npq} + 1/2\right)\left[-z\sqrt{\frac{p}{nq}} - \frac{1}{2}z^2\left(\frac{p}{nq}\right) - \frac{1}{3}z^3\left(\frac{p}{nq}\right)^{3/2} - \cdots\right]$$

$$\log N = \left[\left(np + z\sqrt{npq} + 1/2\right)\left(z\sqrt{\frac{q}{np}}\right) - \left(np + z\sqrt{npq} + 1/2\right)\left(\frac{1}{2}z^2\left(\frac{q}{np}\right)\right)\right.$$

$$\left. + \left(np + z\sqrt{npq} + 1/2\right)\left(\frac{1}{3}z^3\left(\frac{q}{np}\right)^{3/2}\right) - \cdots\right]$$

$$+\left[\left(nq-z\sqrt{npq}+1/2\right)\left(-z\sqrt{\frac{p}{nq}}\right)+\left(nq-z\sqrt{npq}+1/2\right)\left(-\frac{1}{2}z^2\left(\frac{p}{nq}\right)\right)+\left(nq-z\sqrt{npq}+1/2\right)\left(-\frac{1}{3}z^3\left(\frac{p}{nq}\right)^{3/2}\right)-\cdots\right]$$

$$\log N=\left[\left\{z\sqrt{npq}-\frac{1}{2}qz^2+\frac{1}{3}z^3\frac{q^{3/2}}{\sqrt{np}}+z^2q-\frac{1}{2}z^3\frac{q^{3/2}}{\sqrt{np}}+\frac{1}{2}z\sqrt{\frac{q}{np}}-\frac{1}{4}z^2\frac{q}{np}+\ \ldots\right\}\right]+\left[\left\{-z\sqrt{npq}-\frac{1}{2}pz^2-\frac{1}{3}z^3\frac{p^{3/2}}{\sqrt{nq}}+z^2p+\frac{1}{2}z^3\frac{p^{3/2}}{\sqrt{nq}}-\frac{1}{2}z\sqrt{\frac{p}{nq}}-\frac{1}{4}z^2\frac{p}{nq}+\ \ldots\right\}\right]$$

$$\log N=\left[-\frac{1}{2}z^2(p+q)+z^2(p+q)+\frac{z}{2\sqrt{n}}\left(\sqrt{\frac{q}{p}}-\sqrt{\frac{p}{q}}\right)+\theta'(n^{-1/2})\right]$$

$\log N=\left[\frac{1}{2}z^2+\theta(n^{-1/2})\right]\rightarrow\frac{z^2}{2}$ untuk $n\rightarrow\infty$.

Jadi, $\lim_{n\to\infty}\log N=\frac{z^2}{2}$ maka $\lim_{n\to\infty}N=\exp\left(\frac{z^2}{2}\right)$ dan $dG(z)=g(z)dz=\frac{1}{\sqrt{2\pi}}\exp\left(\frac{z^2}{2}\right)$, $-\infty<z<\infty$. Sehingga fungsi probabilitas z adalah $g(z)=\frac{1}{\sqrt{2\pi}}\exp\left(\frac{z^2}{2}\right)$, $-\infty<z<\infty$.

Ini merupakan distribusi normal dengan mean sama dengan nol dan variansi satu. Jika X merupakan variabel acak berdistribusi normal dengan mean μ dan standar deviasi σ maka $Z=\frac{X-\mu}{\sigma}$ adalah distribusi normal baku dan Jacobian adalah $1/\sigma$, sehingga $f(x;\mu,\sigma)=\frac{1}{\sigma\sqrt{2\pi}}exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$; $-\infty<x<\infty$.

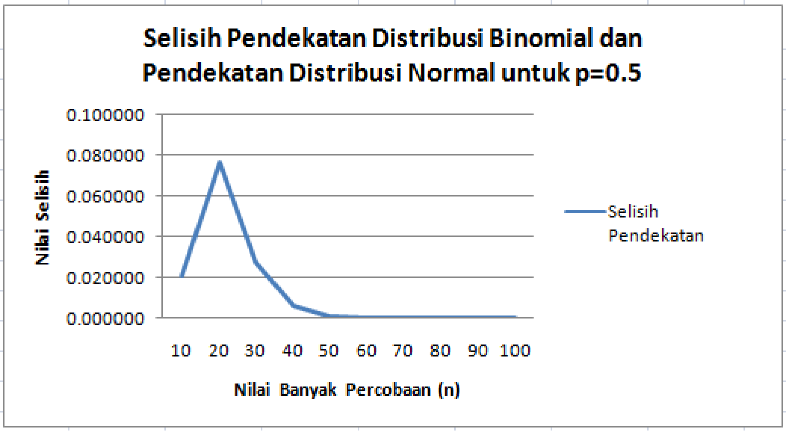
Bentuk $f(x)$ di atas adalah distribusi normal dengan $mean=\mu$ dan standar deviasi $=\sigma$.

Untuk melihat selisih pendekatan distribusi binomial dan pendekatan distribusi normal dapat dilakukan percobaan perhitungan $P(Batas\_Atas\leq X\leq Batas\_bawah)$ dengan menggunakan distribusi binomial untuk $p=0{,}5$ dan berbagai nilai n. Hasilnya dapat lihat pada Tabel 1.

Tabel 1. Perbandingan Antara Distribusi Binomial dan Normal untuk beberapa nilai parameter n dan $p = 0{,}5$

No	n	p	Batas Bawah	Batas Atas	Dengan binomial	Dengan normal	Selisih Pendekatan
1	10	0,5	3	7	0,773438	0,794097	0,020659
2	20	0,5	5	9	0,391207	0,314687	0,076520
3	30	0,5	7	11	0,097633	0,070320	0,027313
4	40	0,5	9	13	0,018899	0,013176	0,005723
5	50	0,5	11	15	0,003255	0,002301	0,000954
6	60	0,5	13	17	0,000527	0,000389	0,000138
7	70	0,5	15	19	0,000082	0,000065	0,000018
8	80	0,5	17	21	0,000013	0,000011	0,000002
9	90	0,5	19	23	0,000002	0,000002	0,000000
10	100	0,5	21	25	0,000000	0,000000	0,000000

Dari Tabel 1 terlihat bahwa semakin besar n, maka selisih perhitungan antara distribusi binomial dan distribusi normal akan mendekati nol. Apabila selisih pendekatan tersebut digambarkan berupa grafik maka terlihat penurunan secara eksponensial dari selisih tersebut apabila $n \to \infty$. Hal ini dapat dilihat pada Gambar 5 di bawah.



Gambar 5. Grafik Selisih Perhitungan Distribusi Binomial dengan Distribusi Normal untuk p=0,5

Pengambilan nilai $p = 0{,}5$ yang tetap, sebab hal ini untuk menunjukkan bahwa untuk kenaikan nilai n terlihat bahwa nilai selisihnya akan semakin mendekati nol. Sedangkan untuk nilai $p \to 0$ dan $p \to 1$, selisih pendekatan itu semakin besar. Alasan logis dari hal ini dapat dilihat secara visual dari Gambar 1.(a) dan Gambar 1.(b) , bahwa distribusi binomial dengan $p \to 1$ berbentuk menceng kiri dan distribusi binomial dengan $p \to 0$ akan menceng kanan. Sehingga dari sifat distribusi binomial yang menceng akan berbeda dengan distribusi normal yang mempunyai sifat simetris.

Aplikasi praktis dari pembuktian ini dapat ditemukan dalam pengolahan big data dan machine learning. Pada dataset yang sangat besar, menghitung probabilitas secara eksplisit menggunakan distribusi binomial sering kali menjadi tidak efisien karena melibatkan perhitungan faktorial dan operasi numerik yang kompleks. Sebagai solusinya, apabila jumlah percobaan n cukup besar dan peluang sukses p tidak terlalu ekstrem (tidak terlalu dekat dengan 0 atau 1), distribusi binomial dapat didekati menggunakan distribusi normal. Pendekatan ini memungkinkan penggunaan metode analitik dan algoritma yang lebih cepat untuk estimasi probabilitas. Dalam aplikasi machine learning, banyak algoritma seperti regresi logistik dan naive Bayes bergantung pada asumsi probabilitas. Untuk variabel biner dalam skala sangat besar, distribusi binomial sering digunakan sebagai dasar komputasi. Namun, dalam situasi berskala besar, distribusi normal sering kali menggantikan binomial untuk mengurangi kompleksitas komputasi.

KESIMPULAN

Berdasarkan penjabaran dan pembahasan di atas, fungsi dari *moment generating function* memiliki peran yang sangat penting dalam Matematika Statistik, khususnya dalam menunjukkan konvergensi distribusi binomial ke distribusi normal melalui kemiripan fungsi pembangkit momennya. Selain itu, rumus Stirling juga memiliki peran krusial dalam menyederhanakan perhitungan faktorial besar pada distribusi binomial, sehingga memungkinkan pendekatan analitik yang lebih sederhana ke distribusi normal. Penelitian ini memberikan manfaat bagi pembaca dalam memahami proses pembuktian secara runtut dan terstruktur, yang dapat diaplikasikan pada kasus serupa. Lebih lanjut, penelitian ini juga membuka peluang untuk pengembangan lebih lanjut, seperti penerapan rumus Stirling pada berbagai kasus lain atau transformasi distribusi lain yang relevan dalam ilmu Statistika.

UCAPAN TERIMA KASIH

Penulis berterima kasih kepada Fakultas Teknologi Informasi UKDW yang telah memberi dukungan dan memfasilitasi penelitian ini dari awal sampai dengan akhir.

REFERENSI

Diana. (2017). Distribusi Binomial sebagai Estimasi Probabilitas Kesuksesan pada Uji Coba Kualitas Layanan Sistem Informasi. *Jurnal Ilmiah MATRIK Vol. 19 No. 3 , Desember 2017*, 227-236.

Hermawan, T., Fajarini, N. m., & Utami, N. (2019). Statistical Reasoning Tentang Kebijakan Satu Anak per Keluarga di Negeri Tirai Bambu (Cina). *Jurnal Intersections Volume 4, No. 2, Agustus 2019 P-ISSN 2685-7952*, 22-32.

Hogg, R. V., McKean, J. W., & Craig, A. T. (2019). *Introduction to Mathematical Statistics Eight Edition.* Boston: Pearson.

Loban, J. M., Sy, Y. P., & Tang, M. I. (2023). Uji Distribusi Binomial pada Lama Studi Mahasiswa. *SEPREN: Journal of Mathematics Education and Applied Vol. 4, No. 2, May 2023 E-ISSN: 2686-4452*, 140-146.

Mann, P. S., & Lacke, C. J. (2021). *Introductory Statistics 10th Edition.* John Wiley & Sons, Inc.

Musadi, & Kurniawati, I. (2023). Normal Distribution Approximation Through Binomial and Poisson Distribution. *Journal of Mathematics and Mathematic EducationVolume 13, No. 1, June 2023 p-ISSN 2089-8878 e-ISSN 2715-8276*, 108-122.

Pebriyanto, A., Sartika, D., Ruspandi, I., Zihani, N., Sam, M., Gifari, M., . . . Pradana, R. (2021). Distribusi Binomial sebagai Pengukuran Keberhasilan dan Kegagalan Produksi Home Industri @One Hand Made. *Buletin of Applied Industrial Engineering Theory, Vol 2 No. 2, Sept 2021 p-ISSN 2720-9628 e-ISSN 2720-961x.*

Setyaningsih, A., Gunawan, M. I., Taher, R. A., & Fauzi, L. (2021). Metode Binomial mengenai Keberhasilan Pemerintah dalam Mengatasi Kemacetan di Ibukota Jakarta. *Buletin of Applied Industrial Engineering Theory , Vol 2 No. 1, Maret 2021 p-ISSN 2720-9628 e-ISSN 2720-961x.*

Sitopu, J. W., & Siswadi. (2022). Fungsi Pembangkit Momen dari Distribusi Probabilitas Diskrit. *FARABI: Jurnal Matematika dan Pendidikan Matematika (JMPM) Volume 5, Nomor 2, Desember 2022, p-ISSN 2623-2332 e-ISSN 2798-5474*, 144-153.

Walck, C. (10 September 2007). *Hand-book on Statistical Distributions for Experimentalist, Particle Physics Group Fysikum Univ of Stockholm.* Stockholm: University of Stockholm.

Walpole, R. E., Myers, R. H., Myers, S. L., & Ye, K. (2012). *Probability & Statistics for Engineers & Scientist Ninth Edition.* Boston: Prentice Hall.