PENDUGAAN SELANG KEPERCAYAAN EKSAK CLOPPER-PEARSON PADA PROPORSI BINOMIAL MENGGUNAKAN *TAIL FUNCTION* DENGAN *SAMPLE SIZE DETERMINATION*

SKRIPSI

NEHEMIA MILLENIUM PAYUNG

H051181325

PROGRAM STUDI STATISTIKA DEPARTEMEN STATISTIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS HASANUDDIN

MAKASSAR

JULI 2023

PENDUGAAN SELANG KEPERCAYAAN EKSAK CLOPPER-PEARSON PADA PROPORSI BINOMIAL MENGGUNAKAN *TAIL FUNCTION* DENGAN *SAMPLE SIZE DETERMINATION*

SKRIPSI

Diajukan sebagai salah satu syarat memperoleh gelar Sarjana Sains pada Program Studi Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin

NEHEMIA MILLENIUM PAYUNG
H051181325

PROGRAM STUDI STATISTIKA DEPARTEMEN STATISTIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS HASANUDDIN

MAKASSAR

JULI 2023

LEMBAR PERNYATAAN KEOTENTIKAN

Saya yang bertanda tangan di bawah ini menyatakan dengan sungguh-sungguh bahwa skripsi yang saya buat dengan judul:

PENDUGAAN SELANG KEPERCAYAAN EKSAK CLOPPER-PEARSON PADA PROPORSI BINOMIAL MENGGUNAKAN *TAIL FUNCTION* DENGAN *SAMPLE SIZE DETERMINATION*

adalah benar hasil karya saya sendiri, bukan hasil plagiat dan belum pernah dipublikasikan dalam bentuk apapun

Makassar, 17 Juli 2023

METERAI TEMPEL

AFFB0AKX605803645

<u>Nehemia Millenium Payung</u>

NIM H051181325

...

PENDUGAAN SELANG KEPERCAYAAN EKSAK CLOPPER-PEARSON PADA PROPORSI BINOMIAL MENGGUNAKAN *TAIL FUNCTION* DENGAN *SAMPLE SIZE DETERMINATION*

Disetujui Oleh:

Pembimbing Utama	Pembimbing Pertama
Sitti Sahriman, S.Si., M.Si.	**Dr. Nurtiti Sunusi, S.Si., M.Si.**
NIP. 19881018 201504 2 002	**NIP. 19720117 199703 2 002**

Ketua Program Studi

Dr. Anna Islamiyati, S.Si., M.Si.

NIP. 19770808 200501 2 002

Pada 17 Juli 2023

HALAMAN PENGESAHAN

Skripsi ini diajukan oleh :

Nama	: Nehemia Millenium Payung
NIM	: H051181325
Program Studi	: Statistika
Judul Skripsi	: Pendugaan Selang Kepercayaan Eksak Clopper-Pearson Pada Proporsi Binomial Menggunakan *Tail Function* Dengan *Sample Size Determination*

Telah berhasil dipertahankan dihadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Sains pada Program Studi Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin.

DEWAN PENGUJI

1. Ketua : Sitti Sahriman, S.Si., M.Si. (..............................)
2. Sekretaris : Dr. Nurtiti Sunusi, S.Si., M.Si. (..............................)
3. Anggota : Andi Kresna Jaya, S.Si., M.Si. (..............................)
4. Anggota : Anisa, S.Si., M.Si (..............................)

Ditetapkan di : Makassar

Tanggal : 17 Juli 2023

KATA PENGANTAR

HALELUYA! Biarlah Segala yang bernafas memuji TUHAN! HALELUYA! Atas Pertolongan dan Perkenaan-Nya lah penulis dapat menyelesaikan tugas akhir ini. Karena TUHANlah yang memberikan hikmat, dari mulut-Nya datang pengetahuan dan kepandaian. Sebab segala sesuatu adalah dari Dia, dan oleh Dia, dan kepada Dia: Bagi Dialah kemuliaan sampai selama-lamanya! *Soli Deo Gloria*.

Penulis mengucapkan terima kasih sekaligus bersyukur untuk kedua orang tua terkasih **Daniel Payung** dan **Yuliana Lolo** yang selalu mendoakan, mendukung, memberikan kasih dengan tulus ikhlas sampai saat ini. Ucapan terima kasih juga penulis sampaikan kepada seluruh **Keluarga Besar** dimana pun berada yang telah mendoakan.

Penghargaan yang tulus dan ucapan terima kasih dengan penuh keikhlasan juga disampaikan kepada:

1. **Bapak Prof. Dr. Ir. Jamaluddin Jompa, M.Sc.**, selaku Rektor Universitas Hasanuddin beserta seluruh jajarannya.
2. **Bapak Dr. Eng. Amiruddin**, selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin beserta seluruh jajarannya.
3. **Ibu Dr. Anna Islamiyati, S.Si., M.Si.**, selaku Ketua Departemen Statistika, segenap Dosen Pengajar dan staf yang telah membekali ilmu dan kemudahan kepada penulis dalam berbagai hal selama menjadi mahasiswa di Departemen Statistika.
4. **Ibu Sitti Sahriman, S.Si., M.Si.**, selaku Pembimbing Utama dan **Ibu Dr. Nurtiti Sunusi, S.Si., M.Si.**, selaku Pembimbing Pertama yang dengan penuh kesabaran telah meluangkan waktu dan pemikirannya di tengah berbagai kesibukan dan prioritasnya untuk senantiasa memberikan arahan, dorongan, dan motivasi kepada penulis mulai dari awal hingga selesainya penulisan tugas akhir ini.

5. **Bapak Andi Kresna Jaya, S.Si., M.Si.**, dan **Ibu Anisa, S.Si., M.Si.**, selaku Tim Penguji yang telah memberikan kritikan yang membangun dalam penyempurnaan penyusunan tugas akhir ini serta waktu yang telah diberikan kepada penulis.
6. **Ibu Sitti Sahriman, S.Si., M.Si.**, selaku Penasehat Akademik penulis. Terima kasih atas segala bantuan serta arahan yang telah diberikan kepada penulis selama menjalani Pendidikan di Departemen Statistika.
7. Teman-teman yang telah membantu penulis dalam menyelesaikan tugas akhir ini, secara khusus **Kevin, Agra, Ikhsan, Ika, Haksar dan Manto**.
8. Teman-teman grup WA **Payung Imut**, yang telah menjadi teman setia bagi penulis dalam masa-masa menyelesaikan tugas akhir ini.
9. Teman-teman **Statistika 2018**, terima kasih untuk kebersamaan yang telah dilewati selama menempuh Pendidikan di Universitas Hasanuddin. Untuk suka dan duka dalam kerja sama serta dukungan yang telah diberikan kepada penulis saat menghadapi kesulitan atau hambatan dalam penyelesaian tugas akhir ini.
10. Semua pihak yang tidak dapat saya sebutkan satu-persatu yang selalu memberikan doa, semangat dan motivasi kepada penulis sehingga tugas akhir ini boleh terselesaikan.

Semoga tugas akhir ini dapat bermanfaat dan berkontribusi untuk penelitian selanjutnya.

Makassar,17 Juli 2023

Nehemia Millenium Payung

PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIK

Sebagai civitas akademik Universitas Hasanuddin, saya yang bertanda tangan di bawah ini:

Nama	: Nehemia Millenium Payung
NIM	: H051181325
Program Studi	: Statistika
Departemen	: Statistika
Fakultas	: Matematika dan Ilmu Pengetahuan Alam
Jenis Karya	: Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Hasanuddin **Hak Bebas Royalti Non-ekslusif** (***Non-exclusive Royalty- Free Right***) atas tugas akhir saya yang berjudul:

"Pendugaan Selang Kepercayaan Eksak Clopper-Pearson Pada Proporsi Binomial Menggunakan *Tail Function* Dengan *Sample Size Determination*"

Beserta perangkat yang ada (jika diperlukan). Terkait dengan hal di atas, maka pihak universitas berhak menyimpan, mengalih-media/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan memublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.
Dibuat di Makassar pada tanggal, 17 Juli 2023

Yang menyatakan

(Nehemia Millenium Payung)

ABSTRAK

Tail function merupakan salah satu metode pembentukan selang kepercayaan yang mana mengspesifikasikan area ekor dari peubah acak pivot dalam bentuk fungsi parameter target. Salah satu keunggulan dari penggunaan *tail function* adalah diperolehnya selang kepercayaan yang lebih sempit saat informasi prior tersedia. Selain dari itu, penggunaan *tail function* tetap mempertahankan sifat *coverage* dari selang kepercayaan eksak. Ukuran lain yang digunakan untuk mengukur performa dari selang kepercayaan selain dari ukuran *coverage probability* adalah ukuran *expected length*-nya. Penelitian ini akan mengilustrasikan penggunaan *tail function* dalam membentuk selang kepercayaan eksak untuk parameter p dari distribusi binomial yang memiliki *expected length* tertentu. Penelitian sebelumnya oleh Puza (2006) memperlihatkan bahwa interval yang diperoleh menggunakan *tail function* lebih sempit dari interval biasa saat $\bar{x}$ berada di sekitar η, yaitu informasi prior dari eksperimen. Namun, ukuran kedekatan $\bar{x}$ terhadap η agar diperoleh interval tersempit tidak diketahui. Dalam penelitian ini diperlihatkan bahwa ukuran kedekatan ini bergantung pada nilai ξ, yaitu nilai toleransi antara *expected length* dengan target Δ yang ingin dicapai dari selang kepercayaannya. Dari keempat kasus yang dibahas, diperoleh bahwa selang kepercayaan tersempit terjadi saat $\bar{x} \in (\eta - \xi, \eta + \xi)$.

Kata Kunci: Selang Kepercayaan, *Tail Function, Coverage Probability,* Proporsi Binomial, *Expected Length*

ABSTRACT

Tail function is one of many methods to construct confidence interval. This method relying on specifying tail area of pivot random variable in terms of the target parameter. One benefit in using this method is having shorter confidence interval when prior information is available. Other than that, tail function also maintains coverage properties of exact confidence interval. Another measure which is used to evaluate the performance of confidence interval other than coverage probability is expected length. This research will illustrate how to use tail function to construct confidence interval for binomial proportion with desired expected length. Recent paper by Puza (2006) shows that the using of tail function makes confidence interval shorter when $\bar{x}$ *is near* η*, which is prior information of the experiment. However, the nearness measure of* $\bar{x}$ *to* η *is still unknown. In this research, we conjecture the nearness of* $\bar{x}$ *to* η *is dependent on* ξ*, which is tolerance value of expected length with its desired target* Δ *of the confidence interval. This conjecture is support by the results of four cases that have been discuss in this research, which is, the interval is shorter when* $\bar{x} \in (\eta - \xi, \eta + \xi)$.

Keywords: *Confidence Interval, Tail Function, Coverage Probability, Binomial Proportion, Expected Length*

DAFTAR ISI

DAFTAR TABEL

DAFTAR GAMBAR

DAFTAR LAMPIRAN

BAB I
PENDAHULUAN

1.1 Latar Belakang

Ekperimen *Bernoulli* adalah sebuah eksperimen acak yang mana hasil dari eksperimennya dapat diklasifikasikan dalam dua cara yang saling eksklusif dan lengkap, contohnya seperti "sukses" dan "gagal" (Hogg, dkk, 2019). Eksperimen *Bernoulli* ini jika dilakukan sebanyak n kali secara independen dengan peluang "sukses" (p) tetap sama dari satu percobaan ke percobaan yang lain akan membentuk distribusi binomial dengan parameter n dan p. Parameter p selanjutnya disebut sebagai proporsi binomial.

Inferensi untuk proporsi binomial p merupakan salah satu masalah statistik yang paling umum ditemui, dengan aplikasi penting dalam berbagai area seperti uji klinik, analisis risiko dan pengendalian kualitas (Thulin, 2014). Inferensi ini biasanya dilakukan dengan melakukan uji hipotesis atau membuat selang kepercayaannya. Uji hipotesis akan menggunakan data observasi untuk menguji suatu hipotesis statistik mengenai suatu parameter. Sedangkan selang kepercayaan akan menggunakan data observasi untuk mengestimasi himpunan nilai dimana parameter mungkin berada dengan peluang tertentu. Selang kepercayaan juga dapat digunakan untuk menguji suatu hipotesis statistik karena pada dasarnya selang kepercayaan adalah himpunan nilai dimana hipotesis *null* diterima.

Saat membuat selang kepercayaan untuk proporsi binomial p, terdapat dua pilihan, yaitu menggunakan selang kepercayaan eksak atau selang kepercayaan aproksimasi. Interval Wald adalah selang kepercayaan aproksimasi yang umum dikenal sebagai interval standar, dengan selang kepercayaannya didasarkan pada uji pendekatan sampel besar untuk kasus binomial. Namun, ukuran *coverage probability* dari interval Wald biasanya berada di bawah taraf nominal $1-\alpha$, yang membuat selang kepercayaan ini tidak dianjurkan penggunaannya (Brown, dkk, 2001). Hal ini bahkan dapat terjadi untuk ukuran sampel yang besar. Beda halnya dengan selang kepercayaan eksak, yang mana penggunaannya tidak memiliki risiko bahwa *coverage probability* berada dibawah $1-\alpha$.

Salah satu selang kepercayaan eksak yang terkenal yaitu interval Clopper-Pearson. Terdapat beberapa alasan mengapa interval Clopper-Pearson adalah selang kepercayaan yang paling sering digunakan. Satu diantaranya adalah kemudahan dalam hal komputasinya. Namun demikian, interval ini sering dianggap terlalu konservatif karena relatif lebih lebar dari interval yang lain. Selang kepercayaan eksak lain yang lebih sempit dari interval Clopper-Pearson diantaranya: interval Blyth-Still-Casella, yang mana dipastikan merupakan selang kepercayaan eksak yang lebih sempit, namun tidak memiliki sifat *nested*; interval Sterne menghasilkan selang kepercayaan yang lebih pendek dari interval Clopper-Pearson, namun dalam beberapa kasus menghasilkan dua interval terpisah daripada satu interval terhubung; interval Blaker yang walaupun lebih lebar dari interval Blyth-Still-Casella, tapi selalu berada di dalam interval Clopper-Pearson dan memiliki sifat *nested*. Namun, Interval Blaker terkadang merupakan gabungan dari interval-interval terpisah dan batas atasnya tidak menurun secara tegas dalam α. Secara kontras, interval Clopper-Pearson memiliki sifat *nested*, himpunan yang selalu terhubung dan memiliki batas yang secara tegas monoton dalam α (Thulin, 2014).

Metode pembentukan selang kepercayaan Clopper-Pearson dikenal sebagai *tail method.* Metode ini melibatkan inversi dari sebuah peluang yang melibatkan peubah acak pivot W sehingga menghasilkan dua statistik sedemikian sehingga parameter p yang dicari berada diantara kedua statistik ini. Teori ini memerlukan spesifikasi dari sebuah tingkat kepercayaan, $1-\alpha$, yang mana "memotong" area dari kedua ekor distribusi W sebanyak $\alpha/2$. Metode lain yang mana melibatkan spesifikasi area ini dalam bentuk fungsi dari p dinamakan *tail function*. Keuntungan utama dari pendekatan ini adalah dapat digunakan untuk menghasilkan selang kepercayaan yang lebih pendek saat informasi prior tersedia (Puza & O'neill, 2006). Oleh karena itu, dengan menerapkan *tail function* pada interval Clopper-Pearson maka diharapkan sifat konsrvatisme dari intervalnya dapat dikurangi.

Selain dari ukuran *coverage probability*, ukuran lain yang digunakan untuk mengukur performa selang kepercayaan untuk proporsi binomial adalah ukuran *expected length* (Cai & DasGupta, 2001). Ukuran ini tidak lain adalah mean dari lebar interval yang dibentuk, yaitu mean selisih statistik batas atas dengan batas bawahnya. Dibandingkan dengan selang kepercayaan aproksimasi, selang kepercayaan eksak memerlukan ukuran sampel yang lebih besar untuk mencapai *expected length* tertentu, sehingga masalah terkait penentuan ukuran sampel menjadi hal yang penting dalam pembentukan selang kepercayaan eksak. Oleh karena itu, penelitian ini akan berfokus pada pembentukan interval Clopper-Pearson menggunakan *tail function* dengan *expected length* tertentu. Metode ini akan diaplikasikan pada data studi kasus yang diperoleh dari hasil simulasi bangkitan *Monte-Carlo*.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, rumusan masalah dalam penelitian ini diantaranya

1. Bagaimana estimasi selang kepercayaan eksak Clopper-Pearson untuk proporsi binomial menggunakan *tail function*?
2. Berapa ukuran sampel minimum yang diperlukan untuk mencapai suatu *expected length* yang diberikan dalam membentuk selang kepercayaan eksak Clopper-Pearson menggunakan *tail function* untuk proporsi binomial?

1.3 Batasan Masalah

Batasan masalah dalam penelitian ini agar penelitian tidak meluas diantaranya

1. *Tail function* yang digunakan dalam penelitian ini adalah *Gaussian tail function* dengan parameter fungsi $\delta = 0, 0.01$ dan $\lambda = 0.1$ serta kasus khusus $\lambda \to 0$.
2. Data dari penelitian berasal dari studi kasus yang dibangkitkan menggunakan metode *Monte-Carlo*.
3. Penentuan *expected length* dari selang kepercayaan yang dibentuk didasarkan pada masalah dalam studi kasusnya.

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah diuraikan, tujuan penelitian ini diantaranya

1. Memperoleh estimasi selang kepercayaan eksak Clopper-Pearson untuk proporsi binomial menggunakan *tail function*.
2. Memperoleh ukuran sampel minimum yang diperlukan untuk mencapai suatu *expected length* yang diberikan dalam membentuk selang kepercayaan eksak Clopper-Pearson menggunakan *tail function* untuk proporsi binomial.

1.5 Manfaat Penelitian

Adapun manfaat dari penelitian ini diantaranya

1. Menambah wawasan dan pengetahuan mengenai estimasi selang kepercayaan eksak Clopper-Pearson menggunakan *tail function* dengan *sample size determination*.
2. Sebagai rujukan dan pengembangan pembelajaran statistika mengenai estimasi selang kepercayaan eksak Clopper-Pearson menggunakan *tail function* dengan *sample size determination.*

BAB II
TINJAUAN PUSTAKA

2.1 Prosedur Selang Kepercayaan Eksak untuk Peubah Acak Diskrit

Misalkan $X_1, X_2, \dots, X_n$ merupakan sampel acak dari peubah acak diskrit X dengan *probability mass function* $P(X = x; \theta)$, $\theta \in \Omega$, dimana Ω adalah interval dari bilangan real. Misalkan $Y = Y(X_1, X_2, \dots, X_n)$ adalah estimator untuk θ dengan *cumulative distribution function* $F_Y(y; \theta)$. Misalkan $\alpha_1 > 0$ dan $\alpha_2 > 0$ diberikan sedemikian sehingga $\alpha = \alpha_1 + \alpha_2$. Pilih $\underline{\theta}$ dan $\overline{\theta}$ sebagai solusi dari persamaan

$$F_Y(y-; \underline{\theta}) = 1 - \alpha_2 \tag{2.1}$$

$$F_Y(y; \overline{\theta}) = \alpha_1 \tag{2.2}$$

Dengan $Y-$ adalah statistik yang mana *support* nya memiliki lag satu nilai dari *support* Y. Sebagai contoh, jika $y_i < y_{i+1}$, yang mana merupakan *support* berurutan dari Y, maka $Y = y_{i+1}$ jika dan hanya jika $Y- = y_i$. Dengan kondisi ini, interval $(\underline{\theta}, \overline{\theta})$ merupakan selang kepercayaan untuk θ dengan *confidence coefficient* paling kurang sebanyak $1 - \alpha$ (Hogg dkk, 2019).

2.2 Generalisasi Teori Klasik

Misalkan sebuah distribusi yang mana bergantung pada θ parameter skalar konstan, dan misalkan diberikan sampel acak dari n observasi $X_1, X_2, \dots, X_n$ dari distribusi tersebut, dengan realisasi $x_1, x_2, \dots, x_n$. Misalkan $\boldsymbol{X}$ dan $\boldsymbol{x}$ masing-masing menotasikan vektor $(X_1, X_2, \dots, X_n)$ dan $(x_1, x_2, \dots, x_n)$. Selanjutnya misalkan $W = g(\boldsymbol{X}, \theta)$ adalah fungsi skalar dari $\boldsymbol{X}$ dan θ yang mana monoton dalam θ, dengan realisasi $w = g(\boldsymbol{x}, \theta)$ dan misalkan distribusi dari W dinotasikan oleh $F_W(w)$, yang kontinu dan tidak bergantung pada θ. Maka $F_W(W)$ memiliki distribusi uniform standar, dan W dinamakan sebagai *pivotal quantity*. Selanjutnya pilih tingkat kepercayaan $1 - \alpha$, dimana $\alpha \in (0,1)$. Maka untuk sembarang nilai dimana θ mungkin berada, adalah benar bahwa

$$1 - \alpha = P[\alpha/2 < F_W(W; \theta) < 1 - \alpha/2] \tag{2.3}$$

Teori klasik dari selang kepercayaan melibatkan manipulasi (3) sehingga menghasilkan

$$1 - \alpha = P[L(\boldsymbol{X}) < \theta < U(\boldsymbol{X})] \tag{2.4}$$

Dimana L dan U adalah dua fungsi dari $\boldsymbol{X}$ yang mana tidak bergantung pada θ. Selang kepercayaan $1 - \alpha$ untuk θ kemudian didefinisikan sebagai (ℓ, u), dimana ℓ dan u adalah solusi dalam θ dari persamaan berikut:

$$F_W(W-; \theta) = 1 - \alpha/2 \tag{2.5}$$

$$F_W(W; \theta) = \alpha/2 \tag{2.6}$$

Sekarang misalkan $\tau(\theta)$, sebuah fungsi yang mana tak menurun sepanjang semua kemungkinan nilai dari θ dan memiliki range dalam interval $[0,1]$. Maka sebagaimana (2.3) benar untuk semua θ, begitu pula

$$1 - \alpha = P[\alpha\tau(\theta) < F_W(W; \theta) < 1 - \alpha + \alpha\tau(\theta)] \tag{2.7}$$

Maka selang kepercayaan $1 - \alpha$ untuk θ adalah (ℓ, u), dimana ℓ dan u adalah solusi dalam θ dari masing-masing persamaan berikut:

$$F_W(w-; \theta) = 1 - \alpha + \alpha\tau(\theta) \tag{2.8}$$

$$F_W(w; \theta) = \alpha\tau(\theta) \tag{2.9}$$

Dengan $\tau(\theta)$ dinamakan sebagai *tail function.* Persamaan (2.8) dan (2.9) dapat diselesaikan dengan menggunakan algoritma Newton-Raphson (Puza & O'neill, 2006).

2.3 Tail Function

Tail function yang digunakan dalam penelitian ini dinakaman *Gaussian* dan memiliki bentuk

$$\tau(p) = \delta + (1 - 2\delta)\Phi\left(\frac{p - \eta}{\lambda}\right) \tag{2.10}$$

dimana $\eta \in (0,1)$, $\lambda > 0$, $0 \leq \delta \leq \frac{1}{2}$ dan $\Phi(z)$ menotasikan fungsi distribusi kumulatif dari distribusi standar normal.

Nilai η harus dispesifikasikan sebagai nilai p yang dianggap paling mungkin, untuk kemudian $\bar{X}$ akan dekat ke η sehingga selang kepercayaan yang dihasilkan lebih sempit dibandingkan dengan selang kepercayaan biasa (Interval Wald dan Clopper-Pearson). Apabila diyakini bahwa p berada di antara p_1 dan p_2, maka pemilihan η yang dapat digunakan adalah $\eta = \frac{p_1+p_2}{2}$. Sedangkan untuk δ, spesifikasi nilai yang kecil pada δ akan menghasilkan interval yang relatif lebih pendek jika $\bar{x}$ dekat dengan η dan sebaliknya. Dengan kata lain, penggunaan δ sangat bergantung pada seberapa kuat informasi prior yang dimiliki.

2.4 Selang Kepercayaan Clopper-Pearson

Interval Clopper-Pearson two-sided untuk proporsi p adalah inversi dari uji binomial *equal-tailed*: interval ini mengandung semua nilai p yang tidak ditolak oleh uji pada taraf signifikan α. Diberikan sebuah observasi dari Y, batas bawahnya diberikan oleh nilai ℓ sedemikian sehingga

$$\sum_{i=y}^{n} \binom{n}{i} \ell^i (1-\ell)^{n-i} = \alpha/2 \tag{2.11}$$

Dan batas atas diberikan oleh nilai u sedemikian sehingga

$$\sum_{i=0}^{y} \binom{n}{i} u^i (1-u)^{n-i} = \alpha/2 \tag{2.12}$$

Misalkan $Beta(t, a, b)$ adalah fungsi densitas dari peubah acak $Beta(a, b)$. Maka

$$\sum_{i=y}^{n} \binom{n}{i} p^i (1-p)^{n-i} = \int_0^p Beta(t, y, n-y+1)dt \tag{2.13}$$

Saat (2.13) disubtitusi ke (2.11) dan (2.12), masalah untuk menemukan ℓ dan u tereduksi menjadi inversi dari dua fungsi distribusi beta. Maka dari itu, titik ujung dari interval Clopper-Pearson diberikan oleh kuantil dari distribusi beta:

$$\big(L(y), U(y)\big) = \left(B\left(\frac{\alpha}{2}, y, n-y+1\right), B(1-\alpha/2, y+1, n-y)\right) \tag{2.14}$$

Saat $Y = 0$, intervalnya adalah $\left(0{,}1 - \left(\frac{\alpha}{2}\right)^{\frac{1}{n}}\right)$ dan saat $Y = n$ intervalnya adalah $\left(\left(\frac{\alpha}{2}\right)^{\frac{1}{n}}, 1\right)$ (Thulin, 2014).

Metodologi *tail function* juga dapat diaplikasikan dalam konteks ini. Diberikan *tail function* $\tau(p)$, interval Clopper-Pearson yang dimodifikasi didefinisikan sebagai (ℓ, u) dimana ℓ dan u adalah solusi dalam p dari persamaan:

$$P(Y \geq y; p) = \alpha\{1 - \tau(p)\}, \;\; P(Y \leq y; p) = \alpha\tau(p) \tag{2.15}$$

Kecuali, $\ell = 0$ jika $y = 0$ dan $u = 1$ jika $y = n$. Persamaan ini dapat diselesaikan menggunakan algoritma Newton-Raphson (Puza & O'neill, 2006).

2.4.1. Kasus $\lambda \to 0$ Untuk Interval Clopper-Pearson

Dengan memilih $\lambda \to 0$, akan memaksimumkan kemungkinan penurunan lebar interval pada $(\bar{x} = \eta)$ untuk δ yang diberikan. Adapun batas untuk $\lambda \to 0$:

$$L(y) = \begin{cases} B(\alpha(1-\delta), y, n-y+1), & B(\alpha(1-\delta), y, n-y+1) < \eta \\ B(\alpha\delta, y, n-y+1), & B(\alpha\delta, y, n-y+1) > \eta \\ \eta, & lainnya \end{cases} \tag{2.16}$$

$$U(y) = \begin{cases} B(1-[\alpha(1-\delta)], y+1, n-y), & B(1-[\alpha(1-\delta)], y+1, n-y) > \eta \\ B(1-\alpha\delta, y+1, n-y), & B(1-\alpha\delta, y+1, n-y) < \eta \\ \eta, & lainnya \end{cases} \tag{2.17}$$

Dengan $B(q, a, b)$ adalah kuantil q dari distribusi beta yang memiliki parameter a dan b.

2.5 Prosedur Eksak untuk memperoleh Sample Size

Coverage Probability dan Expected Length diberikan oleh

$$CP(n,p) = \sum_{i=0}^{n} \binom{n}{i} p^i (1-p)^{n-i} I_{[L(i),U(i)]}(p) \tag{2.18}$$

(dengan $I_{[L(i),U(i)]}(p) = 1$ jika $p \in [L(i), U(i)]$ dan $I_{[L(i),U(i)]}(p) = 0$, lainnya) dan

$$EL(n,p) = \sum_{i=0}^{n} \binom{n}{i} p^i (1-p)^{n-i} \big(U(i) - L(i)\big) \tag{2.19}$$

Misalkan bahwa tujuan dari suatu penelitian adalah untuk mengestimasi p, menggunakan selang kepercayaan $100(1-\alpha)\%$ dengan lebar Δ, yang sebelumnya telah ditentukan oleh peneliti. Untuk menentukan ukuran sampel yang dibutuhkan berdasarkan metode Wald adalah

$$n = \left\lceil \frac{4z_{1-\frac{\alpha}{2}}^2 p_0(1-p_0)}{\Delta^2} \right\rceil \tag{2.20}$$

Adapun penentuan ukuran sampel berdasarkan prosedur eksak EL adalah memperoleh nilai n sedemikian sehingga

$$EL(n,p) = \Delta \tag{2.21}$$

Persamaan ini tidak memiliki bentuk tertutup atau sebuah solusi bilangan bulat. Namun, adalah mungkin untuk mencari bilangan bulat yang meminimumkan $|EL(n,p) - \Delta|$.

Langkah pertama: Diberikan α, Δ dan toleransi ξ, maka ukuran sampel adalah nilai bilangan bulat n yang memenuhi

$$|EL(n,p) - \Delta| \leq \xi \tag{2.22}$$

Nilai ξ dapat dipandang sebagai nilai toleransi (mis, $\xi = 10^{-4}$) yang ditentukan oleh peneliti. Pemilihan parameter ekstra ini tidak mempersulit penentuan ukuran sampel karena hal tersebut hanya sebuah ketidaksesuaian yang dapat diterima antara target Δ dengan $EL(n, p)$. Jika tidak terdapat solusi dari persamaan (2.22), maka nilai ξ dinaikkan.

Langkah kedua: Jika terdapat lebih dari satu solusi $(k > 1)$ pada persamaan (2.22), yang dinotasikan oleh $S_\xi = \{n_1, n_2, \dots, n_k\}$, pilih n berdasarkan salah satu kriteria berikut:

i. $n = \text{argmax}_{l \in S_\xi}\{CP(l, p)\}$
ii. $n = \text{argmin}_{l \in S_\xi}\{CP(l, p)\}$
iii. $n = \min\{S_\xi\}$
iv. $n = \text{argmin}_{l \in S_\xi}\{|CP(l, p) - (1 - \alpha)|\}$

Solusi keseluruhan untuk masalah optimisasi ini selalu dapat ditemukan dengan pencarian yang mendalam, atau dengan kata lain, mencoba secara berturut-turut $n = 1,2,3, \dots$, (Gonçalves dkk, 2012).

2.6 Metode Transformasi Invers Monte Carlo

Misalkan X adalah peubah acak dengan fungsi distribusi F dan $U \sim Unif(0,1)$. Definisikan fungsi F^{-1} sebagai

$$F^{-1}(u) = \inf\{x : F(x) \geq u\}, \qquad 0 \leq u \leq 1 \tag{2.23}$$

Karena F adalah cdf yang berarti F adalah fungsi tak-menurun dan kontinu kanan, sehingga jika $F^{-1}(u) \leq x$ maka $u \leq F(x)$ (A. de La Fortelle, 2020). Dan karena $U \sim Unif(0,1)$ maka $P(U \leq u) = u$, sehingga diperoleh

$$P(X \leq x) = P(F^{-1}(U) \leq x) \rightarrow P\big(U \leq F(x)\big) = F(x) \tag{2.24}$$

Jadi, untuk membangkitkan peubah acak X dengan fungsi distribusi F, ambil $U \sim Unif(0,1)$ dan pilih $X = F^{-1}(u)$ (Rubinstein & Kroese, 2017).

2.7 Metode Newton-Raphson

Metode Newton-Raphson juga dikenal sebagai *tangent method* karena metode ini mengaproksimasi kurva dekat akar dengan garis lurus. Misalkan x_0 adalah aproksimasi awal ke akar dari $\psi(x) = 0$. Maka, $P(x_0, \psi_0)$, dimana $\psi_0 = \psi(x_0)$ adalah titik pada kurva. Kurva di persekitaran akar diaproksimasi dengan garis singgung pada kurva di titik $P(x_0, \psi_0)$. Titik potong dari garis singgung dengan sumbu-x diambil sebagai nilai aproksimasi ke akar. Proses ini diulang hingga iterasi yang diperlukan telah dicapai. Persamaan garis singgung pada kurva $y = \psi(x)$ pada titik $P(x_0, \psi_0)$ diberikan oleh

$$y - \psi(x_0) = (x - x_0)\psi'(x_0) \tag{2.25}$$

dimana $\psi'(x_0)$ adalah kemiringan dari garis singgung pada kurva di titik P. Dengan memilih $y = 0$ dan menyelesaikan persamaan (2.25), diperoleh

$$x = x_0 - \frac{\psi(x_0)}{\psi'(x_0)}, \qquad \psi'(x_0) \neq 0 \tag{2.26}$$

Aproksimasi selanjutnya untuk akar diberikan oleh

$$x_1 = x_0 - \frac{\psi(x_0)}{\psi'(x_0)}, \qquad \psi'(x_0) \neq 0 \tag{2.27}$$

Prosedur tersebut diulangi sehingga diperoleh metode iterasi yang didefinisikan oleh (Iyengar & Jain, 2009)

$$x_{k+1} = x_k - \frac{\psi(x_k)}{\psi'(x_k)}, \qquad \psi'(x_k) \neq 0 \tag{2.28}$$