ANALISIS *QUICK COUNT* METODE *MULTISTAGE RANDOM SAMPLING* DENGAN ESTIMASI KONFIDENSI INTERVAL MENGGUNAKAN METODE BAYES

(Studi Kasus: *Quick Count* Pemilihan Presiden 9 Juli 2014 oleh Lembaga Survei Indonesia)

Skripsi

disusun sebagai salah satu syarat

untuk memperoleh gelar Sarjana Sains

Program Studi Matematika

Oleh

Nur Hidayah

4111412016

JURUSAN MATEMATIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS NEGERI SEMARANG

2016

PERNYATAAN

Saya menyatakan bahwa skripsi ini bebas plagiat, dan apabila di kemudian hari terbukti terdapat plagiat dalam skripsi ini, maka saya bersedia menerima sanksi sesuai ketentuan peraturan perundang-undangan.

Semarang, 26 April 2016

METERAI TEMPEL
361FBADF809281049
6000
ENAM RIBU RUPIAH

Nur Hidayah

NIM 4111412016

PENGESAHAN

Skripsi yang berjudul

Analisis *Quick Count* Metode *Multistage Random Sampling* dengan Estimasi Konfidensi Interval Menggunakan Metode Bayes

Disusun oleh

Nur Hidayah

NIM 4111412016

Telah dipertahankan di hadapan sidang Panitia Ujian Skripsi FMIPA UNNES pada Tanggal 26 April 2016.

Panitia:

Ketua

Prof. Dr. Zaenuri S.E, M.Si, Akt

NIP. 196412231988031001

Sekretaris

Drs. Arief Agoestanto, M.Si.

NIP. 196807221993031005

Ketua Penguji

Dra. Sunarmi, M.Si

NIP. 195506241988032001

Anggota Penguji/Pembimbing 1

Prof. YL Sukestiyarno M.S, Ph.D.

NIP. 195904201984031002

Anggota Penguji/Pembimbing 2

Drs. Arief Agoestanto, M.Si.

NIP. 196807221993031005

MOTTO DAN PERSEMBAHAN

MOTTO

- Nikmat Tuhan kamu manakah yang kamu dustakan? (QS. Ar-Rahman: 55).
- Jika tidak ada bahu untuk bersandar, masih ada lantai untuk bersujud.
- Sesungguhnya sesudah kesulitan itu ada kemudahan (QS. Al-Insyirah: 6).
- Barangsiapa bertakwa pada Allah, maka Allah memberi rezeki dari arah yang tidak disangka-sangka. Barangsiapa yang bertakwa kepada Allah, maka Allah jadikan urusannya menjadi mudah. Barangsiapa yang bertakwa kepada Allah, maka Allah akan menghapus dosa-dosanya dan mendapatkan pahala yang agung (QS. Ath-Thalaq: 2, 3,4).

PERSEMBAHAN

- ***Untuk kedua orang tua tercinta Ibu Solikha dan Abah Makki.***
- ***Untuk kakak dan adikku tersayang, Ahmad Zainuddin S.Si dan M. Nayif Musthofa.***
- ***Untuk sahabatku Sely, Selvi, Rija, dan Novi.***
- ***Untuk teman-teman Matematika Angkatan 2012.***
- ***Untuk keluarga kos Wisma Delima 2 Banaran.***
- ***Untuk Universitas Negeri Semarang (Unnes).***

KATA PENGANTAR

Puji syukur kehadirat Allah SWT yang telah memberikan nikmat dan karunia-Nya serta kemudahan sehingga penulis dapat menyelesaikan skripsi yang berjudul "Analisis *Quick Count* Metode *Multistage Random Sampling* dengan Estimasi Konfidensi Interval Menggunakan Metode Bayes".

Penyusunan skripsi ni dapat diselesaikan berkat kerjasama, bantuan, dorongan dari berbagai pihak. Oleh karena itu penulis mengucapkan terima kasih kepada :

1. Prof. Dr. Fathur Rokhman, M.Hum, Rektor Universitas Negeri Semarang.
2. Prof. Dr. Zaenuri, S.E, M.Si, Akt, Dekan FMIPA Universitas Negeri Semarang.
3. Drs. Arief Agoestanto, M.Si, Ketua Jurusan Matematika FMIPA Universitas Negeri Semarang dan Dosen Pembimbing II yang telah memberikan bimbingan, pengarahan, nasehat, saran, dan dorongan selama penyusunan skripsi ini.
4. Drs. Mashuri, M.Si, Ketua Prodi Matematika FMIPA Universitas Negeri Semarang.
5. Prof. YL Sukestiyarno M.S, Ph.D., selaku Dosen Pembimbing I yang telah memberikan bimbingan, pengarahan, nasehat, saran, dan dorongan selama penyusunan skripsi ini.
6. Dra. Sunarmi, M.Si, selaku Dosen Penguji yang telah memberikan penilaian dan saran dalam perbaikan skripsi ini.

7. Prof. Dr. St. Budi Waluyo, M.Si, selaku Dosen Wali saya sejak Semester 1 hingga sekarang yang telah memberikan bimbingan dan arahan.
8. Staf Dosen Matematika Universitas Negeri Semarang yang telah membekali penulis dengan berbagai ilmu selama mengikuti perkuliahan sampai akhir penulisan skripsi ini.
9. Staf Tata Usaha Universitas Negeri Semarang yang telah banyak membantu penulis selama mengikuti perkuliahan dan penulisan skripsi ini.
10. Ibu dan Abah tercinta, Ibu Solikha dan Abah Makki yang senantiasa memberikan dukungan dan doa yang tiada putusnya.
11. Kakak dan Adik tersayang, Ahmad Zainuddin S.Si dan M. Nayif Musthofa yang selalu memberikan motivasi, semangat, dan doa.
12. Teman-Teman Matematika angkatan 2012 yang berjuang bersama untuk mewujudkan cita-cita.
13. Semua pihak yang tidak dapat disebutkan satu per satu yang telah memberikan bantuan.

Penulis menyadari bahwa dalam penyusunan skripsi ini masih terdapat banyak kekurangan. Oleh karena itu, penulis mengharapkan saran dan kritik yang membangun dari pembaca.

Semarang, 26 April 2016

Penulis

ABSTRAK

Hidayah, Nur. 2016. *Analisis Quick Count Metode Multistage Random Sampling dengan Estimasi Konfidensi Interval Menggunakan Metode Bayes*. Skripsi. Jurusan Matematika FMIPA UNNES. Prof. Dr. YL. Sukestiyarno M.S, Ph.D. dan Drs. Arief Agoestanto, M.Si.

Kata Kunci: Bayes, Konfidensi Interval, *Multistage Random Sampling*, *Quick Count*.

Metode *Multistage Random Sampling* dalam perhitungan cepat (*Quick Count)* merupakan teknik *sampling* yang dikontruksikan dari metode *sampling* acak sederhana yang melalui beberapa tahapan pengambilan sampel secara acak. Analisis statistik yang digunakan adalah metode Inferensi Statistik, yaitu berupa perhitungan proporsi perolehan suara untuk mengetahui penyebaran suara untuk masing-masing kandidat dengan mengestimasi konfidensi interval menggunakan metode Bayes.

Permasalahan yang digunakan adalah : 1) bagaimana perhitungan inferensi statistik mencari proporsi ukuran sampel *quick count* metode *Multistage Random Sampling* mewakili populasi dengan mengestimasi konfidensi interval menggunakan metode Bayes, 2) bagaimana analisis akurasi dan presisi *quick count* Pemilihan Umum Presiden 2014 oleh Lembaga Survei Indonesia. Penulisan skripsi ini dengan tujuan mengetahui perhitungan inferensi statistik dengan mengestimasi konfidensi interval metode Bayes dan mengetahui tingkat akurasi dan presisinya.

Metode penelitian yang digunakan dalam penulisan skripsi ini adalah studi literatur atau kajian pustaka dengan tahap-tahap : 1) pemilihan masalah, 2) merumuskan masalah, 3) studi pustaka, 4) pemecahan masalah terdiri dari identifikasi materi prasyarat dan analisis data sebagai studi kasus, 5) penarikan kesimpulan.

Dengan teknik *Multistage Random Sampling* oleh Lembaga Survei Indonesia pada Pemilu tahun 2014 diperoleh ukuran sampel secara proporsional di tingkat provinsi dan secara acak di tingkat Kabupaten/Kota dari 1.568.217 pemilih tersampel yang tersebar di 3990 TPS sehingga estimasi konfidensi interval dengan metode bayes menghasilkan rentang proporsi pemilih kandidat 1 adalah $0{,}4701395808 < \hat{\theta}_B < 0{,}4702619541$ dan kandidat 2 adalah $0{,}5297380457 < \hat{\theta}_B < 0{,}5298604190$. Dari perhitungan $P_A = \frac{\sum_{i=1}^{n} Y_A}{n}$ diperoleh proporsi kandidat 1 adalah 47,02% dan kandidat 2 adalah 52,98%. Oleh karena urutan perolehan suara untuk setiap kandidat adalah sama antara KPU dan LSI maka LSI pada pemilu 2014 memiliki akurasi yang tinggi dan karena selisih hasil perhitungannya masih terletak dalam batas kesalahan yang ditoleransi maka LSI juga memiliki presisi yang tinggi.

Analisis q*uick count* metode *Multistage Random Sampling* pada intinya untuk menganalisis seberapa akurat dan presisi suatu lembaga dengan teknik analisis inferensi statistika salah satunya dengan metode Bayes pada lembaga peneliti yang terpilih sebagai studi kasus. Pada analisis *quick count* digunakan *software* khusus yang dijalankan oleh

organisasi lembaga peneliti yang praktis dan bersifat rahasia, namun harus lebih mengetahui terlebih dahulu perhitungan secara manualnya.

DAFTAR ISI

DAFTAR SIMBOL

X	: Variabel Random X
μ	: Mean Populasi
$\bar{X}$	: Mean Sampel
σ^2	: Variansi Populasi
S^2	: Variansi Sampel dengan Proporsi Populasi
s^2	: Variansi Sampel dengan Proporsi Sampel
N	: Ukuran Populasi
n	: Ukuran Sampel
P	: Proporsi Populasi
p	: Proporsi Sampel
E	: Tingkat Kesalahan *Sampling* (*Margin of error*)
L	: Batas Bawah (*Lower*)
U	: Batas Atas (*Upper*)
$f(X)$	: fungsi densitas peluang dari variabel random X
$E(X)$	: Nilai Ekspektasi dari Variabel Random X
$Var(X)$	: Variansi dari Variabel Random X
μ_Z	: Mean dari Variavel Random Z
μ_X	: Mean dari Variabel Random X
$\mu_{\bar{X}}$	: Mean dari Rata-rata Variabel Random X
σ_X^2	: Variansi dari Variabel Random X
σ_Z^2	: Variansi dari Variabel Random Z

$V(X)$	: Variansi dari Variabel Random X
$V(\bar{X})$	: Variansi dari Rata-rata Variabel Random X
$V(P)$	: Variansi Proporsi Populasi
$\sim$	: Berdistribusi
$\sum_i^n x_i$	: Penjumlahan Himpunan Anggota x_i ... x_n
θ	: Parameter Proporsi
σ	: Standard Deviasi
a	: Parameter Disribusi Beta
b	: Parameter Distribusi Beta
∂	: Differensial
Z	: Variabel Random Normal Standard
α	: Tingkat Signifikansi
$f(\theta)$	: Distribusi Prior
$f(\theta\|x)$	: Distribusi Posterior
Y_A	: Jumlah Pemilih Pasangan Calon A
se_p	: Standar Error Proporsi
P_A	: Proporsi untuk Pasangan Calon A
l	: Wilayah Pemilihan ke-l
k	: Banyak Wilayah Pemilihan

DAFTAR TABEL

Tabel Halaman

DAFTAR GAMBAR

DAFTAR LAMPIRAN

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Asal mula teknik pengumpulan data dengan perhitungan cepat (*quick count*) berawal dari rentetan peristiwa berupa pemberdayaan suara rakyat melalui *polling*. *Quick count* pertama kali digunakan oleh NAMFREL (*National Citizens Movements For Free Election*) yang memantau pelaksanaan pemilu 1986 di Filipina dimana ada dua kandidat yang bersaing ketat yakni Ferdinand Marcos dan Corazon Aquino. NAMFREL berhasil menemukan berbagai kecurangan dan manipulasi suara serta secara meyakinkan dapat menunjukkan kemenangan Cory Aquino, sekaligus menggagalkan klaim kemenangan Marcos. Kebijakan Marcos yang menganulir kemenangan Cory selanjutnya menjadi dasar pembangkangan sipil dan perlawanan rakyat Filipina dalam bentuk *people power* yang berhasil menggulingkan rezim otoriter Marcos. Sehingga secara tidak langsung *quick count* sebagai bagian dari kontrol terhadap pemilu dan bagian dari upaya untuk menegakkan demokrasi dengan mendorong berlangsungnya pemilu yang jujur dan adil.

Quick count telah diterapkan di Indonesia sejak 1997 oleh LP3ES (Lembaga Pelatihan, Penelitian, Penerangan, Ekonomi, dan Sosial) pada pemilu terakhir rezim Soeharto yang dilakukan secara diam-diam bekerjasama dengan salah satu kekuatan politik. *Quick count* ini cukup berhasil, dengan satu hari setelah pelaksanaan pemilu LP3ES mampu memprediksi hasil pemilu di DKI Jakarta persis sebagaimana hasil perhitungan suara oleh LPU (Lembaga Pemilihan Umum). Tetapi karena pertimbangan keamanan dan politik, hasil tersebut tidak diumumkan pada masyarakat. Pada pemilu 1999, LP3ES dengan *quick count* berhasil pula dalam memprediksi secara tepat urutan partai dan persentase suaranya di Provinsi NTB dan pulau jawa. Selanjutnya pada pemilu 2004, LP3ES kembali membuat *quick count* bekerjasama dengan *National Democratic Institute for International Affairs* (NDI), lembaga internasional dari Amerika yang sudah terbiasa dengan perhitungan cepat. LP3ES-NDI secara akurat berhasil memprediksi pemenang pemilu dari urutan 1 sampai 24.

Sejak dimulai dari Pemilihan Umum Presiden Indonesia 2004 hasil prediksi *quick count* tidak pernah menimbulkan kontroversi karena hasil *quick count* tidak jauh berbeda dengan hasil resmi Pemilu. Kontroversi muncul pertama kali dalam pemilihan Presiden 2014, karena adanya dua kelompok lembaga penelitian yang mengumumkan hasil *quick count* yang berbeda. Perbedaan hasil ini kemudian mendapat amplifikasi politis yang luas karena lembaga penelitian SMRC (Saiful Mujani *Research & Consulting*), Litbang Kompas, Radio Republik Indonesia, Lembaga Survei Indonesia, dan Populi Center menunjukkan keunggulan perolehan suara bagi pasangan Jokowi-Jusuf Kalla sedangkan kelompok lembaga LSN (Lembaga Survei Nasional), Puskaptis, Jaringan Suara Indonesia, dan Indonesia Research Center menunjukkan keunggulan pasangan Prabowo-Hatta dalam perolehan suara mereka.

Perhitungan hasil *quick count* sebenarnya sangat sederhana. *Quick count* dilakukan berdasarkan pada pengamatan langsung di Tempat Pemungutan Suara

(TPS) yang telah dipilih secara acak. Unit analisa *quick count* ini adalah TPS, dengan demikian penarikan sampel tidak dapat dilakukan sebelum daftar TPS atau desa yang akan dipantau tersedia. Kekuatan hasil *quick count* sebenarnya bergantung pada bagaimana teknik penarikan sampel dan ukuran sampel pemilih. Sampel tersebutlah yang akan menentukan suara pemilih yang akan dipakai sebagai dasar prediksi hasil pemilu. Sampel yang ditarik secara benar akan memberikan landasan kuat untuk mewakili karakteristik populasi. Salah satu teknik penarikan sampel yang digunakan dalam *quick count* adalah metode *Multistage Random Sampling*. Metode *Multstage Random Sampling* merupakan teknik *sampling* yang dikontruksikan dari metode *sampling* acak sederhana yang melalui beberapa tahapan pengambilan sampel secara acak. Dengan teknik tersebut dimungkinkan setiap anggota populasi mempunyai peluang yang sama untuk dipilih sebagai sampel, sehingga pengukuran dapat dilakukan dengan hanya melibatkan sedikit sampel. Meski tanpa melibatkan semua anggota populasi hasil survei dapat digeneralisasikan sebagai representasi populasi. Sehingga akan diperoleh berbagai macam informasi statistik yang sangat bermanfaat terutama dalam masalah-masalah yang kompleks. Hasil *quick count* atau proporsi ukuran sampel dikatakan memiliki tingkat akurasi yang tinggi jika lembaga penelitian dapat dengan tepat memprediksi pemenang Pemilu dan struktur (posisi) peringkat partai pemenang pemilu. Sedangkan hasil *quick count* dikatakan memiliki presisi yang tinggi jika memiliki selisih proporsi yang kecil untuk masing-masing kandidat antara hasil KPU dan lembaga peneliti. Untuk mengatahui seberapa besar proporsi ukuran sampel *quick count* mampu mewakili populasi sesungguhnya dapat dilihat pada perhitungan estimasi konfidensi interval dengan diketahui ukuran sampel TPS dan ukuran sampel TPS sukses dari populasi. Kegiatan menghitung estimasi konfidensi interval ini merupakan kegiatan statistik inferensi.

Statistik inferensi merupakan salah satu bidang statistik yang berhubungan dengan analisis data sampai pada peramalan atau penarikan kesimpulan mengenai suatu populasi (Supranto, 1992). Tujuan dari statistik inferensi adalah untuk memperoleh informasi tentang suatu populasi berdasarkan informasi yang diperoleh dari sampel (Supranto, 1992). Statistik inferensi terdiri dari dua macam, estimasi dan uji hipotesis (Walpole dan Myers, 1995). Estimasi dibagi menjadi dua, yaitu estimasi titik dan estimasi interval atau yang biasa disebut interval kepercayaan (konfidensi interval). Konfidensi interval adalah selang nilai-nilai estimasi parameter yang mungkin muncul. Derajat kemungkinan tersebut biasanya dinyatakan dengan tingkat kepercayaan (*Confidence Level*), misalnya 95% atau 99%. Dalam praktiknya, kabanyakan konfidensi interval dinyatakan dalam level 95%. Jika tingkat kepercayaannya tinggi dan menghasilkan interval yang sempit, maka nilai parameter tersebut dikatakan "presisi" (Eriyanto, 1992).

Pada suatu penelitian terkadang diamati karakteristik dari sebuah populasi. Beberapa macam ukuran statistik digunakan untuk mengetahui karakteristik dari populasi, misalnya rataan, varian, median, atau proporsi. Pada inferensi statistik untuk mengestimasi konfidensi interval ingin diperoleh kesimpulan mengenai populasi, meskipun tidak praktis untuk mengamati keseluruhan individu yang menyusun populasi atau tidak mungkin jika populasinya tak hingga. Dengan berbagai keterbatasan dan kendala, tidak dimungkinkan mengamati keseluruhan dari elemen populasi, maka dapat dilakukan langkah alternatif yaitu pendugaan populasi dengan menggunakan sampel yang diambil secara acak dari sebuah populasi.

Pada teori estimasi dapat dilakukan dengan dua metode yaitu metode klasik dan metode Bayes (Walpole dan Myers, 1995). Chandra S. tahun 2011 dalam penelitiannya menyatakan jika diketahui ukuran sampel dan ukuran sampel sukses yang sama kemudian dianalisis menggunakan metode klasik dan metode Bayes akan menghasilkan nilai *Mean Square Error* (MSE) dari estimator Bayes lebih kecil dari pada estimator klasik, walaupun estimator Bayes bukan merupakan estimator bias pada parameter θ dari distribusi Binomial. Sehingga dapat dikatakan bahwa estimator Bayes menghasilkan estimator yang baik untuk parameter θ jika MSE dari estimator sebagai ukuran kebaikannya. Hal ini memicu penulis untuk mencari nilai estimasi konfidensi interval pada *quick count* menggunakan metode Bayes. Metode Bayes memandang parameter sebagai variabel yang menggambarkan pengetahuan awal tentang parameter sebelum pengamatan dilakukan dan dinyatakan dalam suatu distribusi yang disebut dengan distribusi prior dikombinasikan dengan informasi dengan data sampel melalui teorema Bayes, dan hasilnya dinyatakan dalam bentuk distribusi yang disebut distribusi posterior yang selanjutnya menjadi dasar untuk inferensi di dalam metode Bayes (Berger, 1990).

Dalam statistik klasik parameter proporsi Binomial dianggap sebagai sebuah nilai yang dianggap konstan, tapi dalam beberapa situasi dan tempat pengamatan yang berbeda akan diperoleh proporsi yang berubah-ubah, sehingga dalam hal ini prinsip Bayes cukup relavan digunakan, karena prinsip Bayes parameter proporsi diperlakukan sebagai variabel agar mempunyai kemampuan yang akomodatif pada keadaan tersebut.

Teorema Bayes memungkinkan seseorang untuk memperbaruhi keyakinannya mengenai sebuah parameter setelah data diperoleh. Sehingga dalam hal ini mengharuskan adanya keyakinan awal (prior) sebelum memulai inferensi. Pada dasarnya distribusi prior bisa diperoleh berdasarkan keyakinan subyektif dari peneliti itu sendiri mengenai nilai yang mungkin untuk parameter yang diestimasi, sehingga perlu diperhatikan bagaimana cara menentukan prior. Jika distribusi sampel berasal dari keluarga eksponensial, maka salah satu caranya adalah dengan menggunakan prior konjugat (Bolstad, 2007), dimana distribusi prior konjugat mengacu pada acuan analisis model terutama dalam pembentukan fungsi *likelihood*nya, sehingga dalam penentuan prior konjugat selalu dipikirkan mengenai penentuan pola distribusi prior yang mempunyai bentuk konjugat dengan fungsi densitas peluang pembangun *likelihood*nya (Box dan Tiao, 1973). Kemudian digabungkan dengan informasi sampel melalui teorema Bayes sehingga dihasilkan distribusi posterior. Setelah distribusi posterior terbentuk, maka dapat diperoleh estimasi titik, interval, dan uji hipotesis Bayes untuk parameter yang diestimasi.

Studi kasus dalam laporan akhir ini adalah *quick count* Pemilihan Umum Presiden 9 Juli 2014 oleh Lembaga Survei Indonesia, lembaga peneliti ini akan dianalisis tingkat akurasi dan presisinya dengan teknik pengambilan sampel menggunakan metode *Multistage Random Sampling* dan estimasi konfidensi interval menggunakan Metode Bayes.

1.2 Rumusan Masalah

Dari uraian di atas diperoleh rumusan masalah dalam tulisan ini adalah sebagai berikut.

1. Bagaimana perhitungan inferensi statistik mencari proporsi ukuran sampel *quick count* metode *Multistage Random Sampling* mewakili populasi dengan mengestimasi konfidensi interval menggunakan metode Bayes ?
2. Bagaimana analisis akurasi dan presisi *quick count* metode *Multistage Random Sampling* dengan estimasi konfidensi interval menggunakan metode Bayes pada *quick count* Pemilihan Umum Presiden 2014 oleh Lembaga Survei Indonesia jika dibandingkan dengan perolehan resmi Komisi Pemilihan Umum ?

1.3 Pembatasan Masalah

Cakupan permasalahan yang disampaikan peneliti dibatasi, yaitu estimasi konfidensi interval menggunakan metode Bayes dan metode Bayes estimator yang digunakan adalah estimator Bayes dengan prior beta serta lembaga peneliti *quick count* Pemilihan Umum Presiden 2014 yang dianalisis adalah Lembaga Survei Indonesia.

1.4 Tujuan Penelitian

Tujuan yang ingin dicapai dari penelitian ini adalah yaitu sebagai berikut.

1. Mengetahui perhitungan inferensi statistik mencari proporsi ukuran sampel *quick count* metode *Multistage Random Sampling* mewakili populasi dengan mengestimasi konfidensi interval menggunakan metode Bayes.
2. Mengetahui analisis akurasi dan presisi *quick count* metode *Multistage Random Sampling* dengan estimasi konfidensi interval menggunakan metode Bayes pada *quick count* Pemilihan Umum Presiden 2014 oleh Lembaga Survei Indonesia jika dibandingkan dengan perolehan resmi Komisi Pemilihan Umum (KPU).

1.5 Manfaat Penelitian

1. Dapat mengetahui teknik *sampling* yang digunakan dalam perhitungan cepat (*quick count*).
2. Dapat mengetahui perhitungan ukuran sampel TPS dalam perhitungan cepat (*quick count*).
3. Dapat mengetahui cara menghitung perkiraan proporsi dengan mengestimasi konfidensi interval menggunakan metode Bayes.
4. Dapat mengetahui analisis akurasi dan presisi *quick count* metode *Multistage Random Sampling* dengan estimasi konfidensi interval menggunakan metode Bayes pada *quick count* Pemilihan Umum Presiden 2014 oleh Lembaga Survei Indonesia jika dibandingkan dengan perolehan resmi Komisi Pemilihan Umum (KPU).

1.6 Sistematika Penulisan

Secara garis besar skripsi ini dibagi menjadi tiga baian yaitu bagian awal skripsi, bagian isi skripsi, dan bagian akhir skripsi. Berikut ini dijelaskan masing-masing bagian skripsi.

1. Bagian awal skripsi

Bagian awal skripsi meliputi halaman judul, abstrak, halaman pengesahan, motto dan persembahan, kata pengantar, daftar isi, daftar simbol, daftar gambar, daftar tabel, dan daftar lampiran.

2. Bagian isi skripsi

Bagian isi skripsi secara garis besar terdiri dari lima bab, yaitu.

BAB I. PENDAHULUAN

Dalam bab ini dikemukakan latar belakang, permasalahan, tujuan penelitian, manfaat penelitian, dan sistematika penulisan skripsi.

BAB II. LANDASAN TEORI

Dalam bab ini dikemukakan konsep-konsep yang dijadikan landasan teori meliputi : konsep dasar survei sampel, teknik *sampling*, *quick count*, cara pemilihan elemen sampel, distribusi binomial, distribusi normal, theorema limit pusat, hubungan antara distribusi normal dan distribusi binomial, perkiraan proporsi, interval keyakinan untuk parameter statistika, interval konfidensi bayes prior beta, uji hipotesis bayes, *margin* *of error*, tingkat kepercayaan, analisis *quick count*, organisasi *quick count*, komunikasi data *quick count*, penelitian terdahulu, dan kerangka berpikir.

BAB III. METODE PENELITIAN

Dalam bab ini dikemukakan metode penelitian yang meliputi : pemilihan masalah, merumuskan masalah, studi pustaka, pemecahan masalah, dan penarikan kesimpulan.

BAB IV. PEMBAHASAN

Dalam bab ini dikemukakan pembahasan mengenai analisis *quick count* metode *multistage random sampling* dengan estimasi konfidensi interval menggunakan metode bayes.

BAB V. PENUTUP

Dalam bab ini dikemukakan simpulan dari pembahasan dan saran yang berkaitan dengan simpulan.

3. Bagian akhir skripsi

Bagian akhir skripsi meliputi daftar pustaka dan lampiran-lampiran yang mendukung.

BAB II

TINJAUAN PUSTAKA

2.1 Landasan Teori

2.1.1 Konsep Dasar Survei Sampel

Dalam suatu penelitian survei, keberadaan populasi dan sampel penelitian tidak dapat dihindarkan. Populasi dan sampel merupakan sumber utama untuk memperoleh data yang dibutuhkan dalam mengungkapkan fenomena atau realitas yang dijadikan fokus penelitian. Demi mencapai keakuratan dan validitas data yang dihasilkan, populasi dan sampel yang dijadikan objek penelitian harus memiliki kejelasan baik dari segi ukuran maupun karakteristiknya. Dengan kata lain, kejelasan populasi dan ketepatan pengambilan sampel dalam penelitian akan menentukan validitas proses dan hasil penelitian. Penjelasan mengenai konsep dasar dalam survei sampel adalah sebagai berikut.

2.1.1.1 Populasi

Populasi atau sering juga disebut *universe* adalah keseluruhan atau totalitas objek yang diteliti yang ciri-cirinya akan diduga atau ditaksir (*estimate*). Ciri-ciri populasi disebut parameter. Oleh karena itu, populasi juga sering diartikan sebagai totalitas semua nilai yang mungkin, hasil menghitung ataupun pengukuran, kuantitatif maupun kualitatif mengenai karakteritik tertentu dari semua anggota kumpulan yang lengkap dan jelas yang ingin dipelajari sifat-sifatnya (Sudjana, 2005). Menurut Margono (2004), populasi diartikan sebagai wilayah generalisasi yang terdiri dari subyek/obyek yang mempunyai kualitas dan karakteristik tertentu yang ditetapkan oleh peneliti untuk dipelajari dan kemudian ditarik kesimpulan.

Konsep dasar dalam populasi yang perlu dipahami adalah jumlah populasi dan ukuran populasi. Jumlah populasi (*population numbers*) adalah banyaknya kategori populasi yang dijadikan objek penelitian. Sedangkan ukuran populasi (*population size*) adalah banyaknya unsur atau unit yang terkandung dalam sebuah kategori populasi tertentu (Kurnia, 1992).

Masalah yang akan muncul dalam pengambilan data berdasarkan seluruh responden populasi adalah masalah biaya, tenaga, dan waktu. Sehingga cenderung peneliti mengambil sampel anggota populasi yang mewakili secara representatif terhadap populasi, dalam hal ini disebut *sampling*.

2.1.1.2 Sensus dan Sampel

Jika peneliti menggunakan seluruh unsur populasi sebagai sumber data, maka penelitiannya disebut sensus. Sensus merupakan penelitian yang dianggap dapat mengungkapkan ciri-ciri populasi secara akurat dan komprehensif, karena dengan menggunakan seluruh unsur populasi sebagai sumber data maka gambaran tentang populasi dapat secara utuh dan menyeluruh akan diperoleh. Jika keadaan peneliti tidak memungkinkan untuk melakukan sensus, maka peneliti dapat mengambil sebagian dari unsur populasi untuk dijadikan objek penelitiannya. Sebagian unsur populasi yang dijadikan objek penelitian disebut sampel. Sampel merupakan bagian dari populasi sedemikian sehingga dapat mewakili atau menggambarkan populasi. Dalam satu populasi dapat mempunyai satu atau lebih sampel, tergantung pada karakteristik dan variabilitas data (Sudjana, 2005).

Alasan-alasan penelitian dilakukan dengan menggunakan sampel adalah sebagai berikut (Supranto, 1992).

a. Ukuran Populasi

Dalam hal populasi tak terbatas (tak terhingga) berupa parameter yang jumlahnya tidak diketahui dengan pasti, pada dasarnya bersifat konseptual. Demikian juga dalam populasi yang terbatas (terhingga) yang jumlahnya sangat besar, tidak praktis untuk mengumpulkan data dari populasi yang jumlahnya sangat besar.

b. Masalah Biaya

Besar-kecilnya biaya tergantung dari banyak-sedikitnya objek yang diselidiki. Semakin besar jumlah objek, maka semakin besar biaya yang diperlukan, lebih-lebih bila objek itu tersebar di wilayah yang cukup luas (seluruh wilayah Indonesia misalnya). Oleh karena itu, penarikan sampel merupakan salah satu cara mengurangi anggaran biaya.

c. Masalah Waktu

Penarikan sampel selalu memerlukan waktu yang lebih sedikit dari pada penelitian menggunakan seluruh populasi. Oleh karena itu, jika waktu penelitian yang tersedia terbatas dan kesimpulan yang diinginkan harus dikumpulkan segera, maka penelitian menggunakan sampel merupakan cara yang sangat tepat untuk lebih mengefisienkan waktu.

Jika peneliti menggunakan sampel sebagai sumber data, maka yang akan diperoleh adalah ciri-ciri sampel bukan ciri-ciri populasi, tetapi ciri-ciri sampel harus dapat digunakan untuk menaksir populasi. Ciri-ciri sampel disebut statistik. Sama halnya dengan populasi, dalam sampelpun ada konsep jumlah sampel dan ukuran sampel. Jumlah sampel adalah banyaknya kategori sampel yang diteliti. Sedangkan ukuran sampel adalah besarnya unsur populasi yang dijadikan sampel. Alasan peneliti harus benar-benar memahami pengertian istilah jumlah sampel dan ukuran sampel adalah karena jumlah sampel dan sifat sampel yang diteliti akan sangat menentukan uji statistik inferensial yang harus digunakan untuk menguji hipotesis yang dirumuskan dalam penelitian (Kurnia, 2015).

Karena data yang diperoleh dari sampel harus dapat digunakan untuk menaksir populasi, maka dalam mengambil sampel dari populasi tertentu peneliti harus benar-benar bisa mengambil sampel yang dapat mewakili populasinya atau disebut sample representatif. Sampel representatif adalah sampel yang memiliki ciri karakteristik yang sama atau relatif sama dengan ciri karakteristik populasinya. Tingkat kerepresentatifan sampel yang diambil dari populasi tertentu sangat tergantung pada jenis sampel yang digunakan, ukuran sampel yang diambil, dan cara pengambilannya. Cara atau prosedur yang digunakan untuk mengambil sampel dari populasi tertentu disebut teknik *sampling* (Kurnia, 2015).

2.1.1.3 Unit Sampling

Unit *sampling* adalah satuan yang didefinisikan untuk pemilihan suatu sampel. Unit *sampling* dapat terdiri atas satu atau lebih unit dasar (Estok *et al*, 2002). Dalam hal penarikan sampel statistik, unit sampel ditetapkan dengan menggunakan formula statistik sesuai dengan jenis *sampling* yang dilakukan. Pada tahap unit *sampling* ini hasilnya berupa pernyataan mengenai jumlah unit sampel yang harus diuji pada populasi yang menjadi objek penelitian.

2.1.1.4 Kerangka (Frame) Sampling

Tingkat kepresentatifan sampel selain ditentukan oleh ukuran sampel yang diambil juga ditentukan oleh teknik *sampling* yang digunakan. Diantara sekian banyak teknik *sampling*, dalam penggunaannya mempersyaratkan tersedianya kerangka *sampling*. Kerangka *sampling* merupakan kumpulan unit *sampling* dan mewakili populasi (Estok *et al*, 2002).

2.1.2 Teknik *Sampling*

2.1.2.1 Sampel Stratifikasi (Stratified Random Sampling)

Sampel stratifikasi (*stratified random sampling*) merupakan teknik penarikan sampel dengan *sampling unit* dikelompokkan menjadi beberapa strata (kelompok) sehingga sampling *unit* dalam satu strata relatif homogen (Scheaffer *et al*, 1990). Adapun alasan digunakan sampel stratifikasi adalah

1. Kesederhanaan dari *simple random sampling*, potensial memperoleh signifikan dalam reabilitas.
2. Populasi harus dibagai dalam k strata yang saling bebas satu sama lain.
3. Penarikan sampel dilakukan secara bebas di setiap strata.

Penetapan ukuran sampel per strata ditentukan oleh tiga faktor berikut.

1. Ukuran populasi setiap strata
2. Ragam setiap strata
3. Biaya pengambilan sampel per strata

Kelebihan dari sampel stratifikasi ini adalah pada waktu melakukan analisis dapat disajikan secara keseluruhan, per strata ataupun membandingkan antar strata.

2.1.2.2 Sampel Klaster (Cluster Sampling)

Sampel klaster (*Cluster sampling*) adalah sampel peluang dengan masing-masing unit sampel (*sampling unit*) merupakan kumpulan atau klaster dari elemen (Scheaffer *et al*, 1990). Elemen didefinisikan sebagai obyek dimana pengukuran akan dilakukan. Sedangkan *sampling unit* mempunyai arti yang hampir sama dengan elemen tetapi ada syarat tidak boleh tumpang tindih. Teknik penarikan sampel pada dasarnya dibedakan menjadi dua yakni berdasarkan kerangka sampel (*sampling frame*) dan tidak berdasar kerangka sampel. *Sampling frame* adalah daftar dari keseluruhan elemen populasi. Teknik berdasarkan kerangka sampel disebut *probabilistic sampling*, dengan memiliki karakteristik setiap elemennya diketahui sehingga penduga tak bias dapat dibuktikan. Sedangkan teknik penarikan sampel tanpa kerangka sampel disebut *nonprobabilistic sampling/ quota/purposive/judgemenet*, teknik ini sering digunakan untuk survei pemasaran dan opini publik.

Cara pengambilan sampel pada *cluster sampling* adalah

1. Populasi dibagi menjadi c klaster
2. Dari c klaster selanjutnya dipilih secara acak sebanyak k klaster
3. Seluruh elemen dari k klaster terpilih diambil.

Sampel klaster merupakan desain yang efektif untuk memperoleh sejumlah informasi khusus dengan biaya minimum bila memenuhi kondisi (Scheaffer *et al*, 1990)

1. *Frame listing* elemen populasi yang baik tidak ada atau sangat mahal, sementara *frame listing* klaster mudah diperoleh.
2. Biaya untuk memperoleh objek-objek yang terpilih sangat mahal karena faktor geografi maka klaster akan mengurangi biaya.

2.1.2.3 Sampling Acak Sederhana

Apabila suatu sampel dengan n elemen dipilih dari suatu populasi dengan N elemen sedemikian rupa sehingga setiap kemungkinan sampel dengan n elemen mempunyai kesempatan yang sama untuk terpilih, maka prosedur *sampling* demikian disebut *sampling* acak sederhana. Peluang yang dimiliki oleh setiap unit penelitian untuk dipilih sebagai sampel sebesar n/N, yaitu ukuran sampel yang dikehendaki dibagi dengan ukuran populasi (Supranto, 1992).

Pemakaian metode *sampling* acak sederhana perlu memenuhi tersedianya kerangka sampel, ukuran populasinya diketahui dengan pasti, dan keadaan populasi tidak terlalu tersebar secara geografis (Eriyanto, 1990).

2.1.2.4 *Multistage Random Sampling*

Metode *Multistage Random Sampling* merupakan teknik *sampling* yang dikontruksikan dari metode *sampling* acak sederhana yang melalui beberapa tahapan pengambilan sampel secara acak. Dengan teknik tersebut dimungkinkan setiap anggota populasi mempunyai peluang yang sama untuk dipilih sebagai sampel, sehingga pengukuran dapat dilakukan dengan hanya melibatkan sedikit sampel. Meski tanpa melibatkan semua anggota populasi, hasil survei dapat digeneralisasikan sebagai representasi populasi. Sehingga akan diperoleh berbagai macam informasi statistik yang sangat bermanfaat terutama dalam masalah-masalah yang kompleks. Ukuran sampel TPS yang digunakan adalah proporsional di masing-masing wilayah pemilihan (LSI, 2006).

Multistage random sampling pada dasarnya adalah gabungan antara sampel stratifikasi (*stratified random sampling*) dengan sampel klaster (*cluster sampling*). Stratifikasi diperlukan agar heterogenitas dari populasi bisa tercermin dalam sampel. Untuk menanggulangi masalah biaya yang meningkat karena stratifikasi tersebut, maka stratifikasi tersebut dikombinasikan dengan klaster. Lewat klaster sampel tidak menyebar sehingga biaya untuk menjangkaunya mengecil meskipun klaster membuat sampel menjadi kurang mencerminkan karaketeristik populasi (Scheaffer *et al*, 1990).

2.1.3 *Quick Count*

Quick count adalah perhitungan cepat hasil pemilihan dengan menggunakan sampel TPS (Tempat Pemungutasn Suara). *Quick count* merupakan kegiatan pengambilan sampel seperti survei yang sering dilakukan untuk mengkaji objek studi tertentu, perbedaannya hanya pada unit terkecil yang diambil dalam sampel. Jika survei unit terkecil adalah desa/kelurahan sedangkan *quick count* adalah TPS. Dengan *quick count*, hasil perhitungan suara bisa diketahui dua sampai tiga jam setelah perhitungan suara di TPS ditutup. Kecepatan perhitungan suara tersebut bisa didapat karena dalam *quick count* tidak menghitung suara dari semua TPS, cukup dengan sampel TPS saja (LSI, 2006).

Karena *quick count* bekerja pada sampel, bekerja pada ketidakpastian data, bekerja pada unit-unit statistik, dan bekerja pada bagian dari populasi, bukan keseluruhan populasi, sehingga ada diskorsi dalam angka yang dihasilkan. Diskorsi adalah gap atau perbedaan atau lebih dikenal dengan *margin of error*. *Margin of error* timbul akibat pengambilan *sampling*. Idealnya sampel akurat adalah sampel yang dihasilkan dari proses *sampling* yang menghasilkan *margin of error* yang kecil atau yang mendekati parameter sesungguhnya dalam populasi. Jika penarikan sampel dilakukan dengan benar, prosedur pencatatan dilakukan dengan tepat, meski hanya memakai sampel TPS, hasil *quick count* akan menggambarkan hasil pemilu.

Quick count yang sukses dimulai dari pemahaman dasar dan tujuan yang jelas. Lembaga peneliti harus dapat mengidentifikasi tujuan-tujuan mereka sehingga memudahkan perencanaan strategi dan taktik pelaksanaannya.

Quick count juga memiliki kemampuan untuk (LSI, 2006)

1. Memberikan indikasi atau dugaan adanaya kecurangan dalam perhitungan suara.

Walaupun pada kasus-kasus tertentu *quick count* tidak dapat mencegah kecurangan, setidaknya data *quick count* dapat memberikan indikasi atau dugaan terjadinya kecurangan dalam perhitungan suara. Hal ini dilakukan dengan mengamati ada tidaknya inkonsistensi perolehan suara di TPS-TPS yang diamati dengan hasil resmi penyelenggara pemilihan. Seringkali kecurangan terungkap ketika hasil tabulasi resmi penyelenggara pemilu berbeda dengan hasil *quick count*.

2. Memprediksi hasil pemilihan secara cepat

Perhitungan perolehan suara resmi oleh penyelenggara pemilihan sering kali memakan waktu lama, sehingga tidak dapat segera diumumkan kepada publik. Lambannya proses ini dapat membuka peluang terjadi nya ketidakpastian atau kekosongan politik yang dapat mengancam stabilitas nasional suatu negara. *Quick count* yang akurat dan kredibel dapat memprediksi secara cepat sehingga mengurangi ketegangan politik setelah pemungutan suara dilakukan. *Quick count* juga dapat meningkatkan kepercayaan warga negara terhadap hasil Pemilu.

3. Melaporkan kualitas Pemilu

Quick count dirancang untuk mengumpulkan informasi secara sistematis dan terpercaya mengenai kualitas Pemilu. Pemantau Independen dapat mengandalkan metode statistik yang digunakan dalam *quick count* untuk memberikan bukti-bukti yang dapat dipercaya mengenai proses pemilu.

2.1.4 Cara Pemilihan Elemen Sampel

1. Dengan cara lotere (pengundian) (Supranto, 1992)

Cara pengundian ini merupakan cara yang paling sederhana dalam memilih sampel. Hal yang perlu dilakukan adalah membuat kerangka sampel yang terdiri dari seluruh elemen populasi, kemudian masing-masing elemen diberikan nomor, dengan syarat setiap elemen mendapat satu nomor. Tahapan selanjutnya nomor-nomor dari seluruh elemen dipilih secara acak, nomor yang terpilih mewakili elemen dari populasi akan menjadi sampel. Cara pengambilan sampel ini hanya cocok diterapkan jika elemen dari populasi jumlahnya sedikit.

2. Dengan menggunakan tabel bilangan acak (Supranto, 1992)

Cara ini meringankan pekerja terutama untuk sampel dari populasi yang besar. Selain itu, juga memberikan jaminan ketelitian yang jauh lebih besar bahwa setiap elemen mempunyai probabilitas yang sama untuk dipilih sebagai sampel. Tabel tersebut mempunyai kolom-kolom yang berisi nomor-nomor dengan 5 angka dan harus ditentukan 5 angka mana yang akan digunakan. Nomor-nomor tersebut dapat dipilih dari titik manapun: kiri atas, kiri bawah, kanan atas atau kanan bawah. Bila telah sampai pada bagian bawah kolom, maka kembali dilanjutkan pada kolom sebelahnya. Hal ini dilakukan terus menerus sampai seluruh sampel diperoleh.

2.1.5 Distribusi Binomial

Suatu percobaan sering terdiri dari beberapa usaha, tiap usaha dengan dua kemungkinan hasil, dapat diberi nama sukses dan gagal. Percobaan tersebut dinamakan percobaan Binomial jika (Walpole dan Myers, 1995)

1. percobaan terdiri atas n usaha yang berulang
2. tiap usaha memberikan hasil yang dapat ditentukan dengan sukses atau gagal
3. peluang sukses dinyatakan dengan θ, tidak berubah dari usaha yang satu ke usaha yang berikutnya
4. tiap usaha bebas dari usaha lainnya.

Distribusi binomial dapat dipandang sebagai n variabel *random* Bernaulli yang *independent*, yaitu banyaknya yang sukses dalam n *trial* bernaulli. Misalkan X variabel *random* yang didefinisikan seperti pada percobaan Bernaulli, jika peluang sukses p dan peluang gagal $q = 1 - p$, maka fungsi peluangnya

$$f(X) = b(X; n, p) = \begin{cases} \binom{n}{x} p^x q^{n-x}, & x = 0,1,2, \dots, n \\ 0 & lainnya. \end{cases} \quad (2.1)$$

(Walpole dan Myers, 1995).

2.1.5.1 *Nilai Harapan*

Nilai harapan suatu variabel random X adalah

$$\mu = E(X) = \sum_x x f(x) \text{ bila X diskrit}$$

$$\mu = E(X) = \int_{-\infty}^{\infty} x f(x) dx \text{ bila X kontinu} \quad (2.2)$$

(Walpole dan Myers, 1995).

Variabel acak yang mampu menjalanai bilangan bulat adalah variabel acak diskrit, sedangkan variabel acak yang mampu menjalani bilangan *real* adalah variabel acak kontinu.

Maka, untuk X variabel random binomial dengan fungsi massa peluang

$$f(X) = b(X; n, p) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x}, & x = 0,1,2,\dots,n \\ 0 & lainnya. \end{cases}$$

$$E(X) = \sum_{i=1}^{n} x_i f(x_i)$$

$$= \sum_{x=0}^{n} x \binom{n}{x} p^x (1-p)^{n-x} = \sum_{x=1}^{n} x \binom{n}{x} p^x (1-p)^{n-x}$$

$$= \sum_{x=1}^{n} n \binom{n-1}{x-1} p^x (1-p)^{n-x} \qquad (2.3)$$

(Walpole dan Myers, 1995).

Jika $y = x - 1$

$$= \sum_{y=1}^{n} n \binom{n-1}{y} p^{y-1} (1-p)^{n-y-1}$$

$$= np \sum_{y=0}^{n-1} n \binom{n-1}{y} p^y (1-p)^{n-y-1} = np \qquad (2.4)$$

(Walpole dan Myers, 1995).

Catatan :

$$\binom{n}{x} = \frac{n!}{(n-x)!\,x!}$$

$$x \binom{n}{x} = \frac{x.\,n!}{(n-x)!\,x!}$$

$$= \frac{n(n-1)!}{(n-x)(x-1)!}$$

$$= n\binom{n-1}{x-1}$$

Dengan menggunakan teorema binomial

$$(a-b)^n = \sum_{x=0}^{n} \binom{n}{x} a^x b^{n-x}$$

Maka, persamaan (2.4) menjadi

$$\sum_{y=0}^{n-1} \binom{n-1}{y} p^y (1-p)^{(n-y)-1} = [p + (1-p)]^{(n-1)}$$

$$= 1^{(n-1)} = 1 \qquad (2.5)$$

(Walpole dan Myers, 1995).

2.1.5.2 *Variansi*

Variansi dari suatu variabel random X, ditulis Var(X) = σ_X^2.

$$Var(X) = E[(X - E(X))^2] = E(X^2) - [E(X)]^2$$

$$E(X^2) = \begin{cases} \sum_{i=1}^{n} x_i^2 f(x_i), untuk\ X\ diskrit \\ \int_{-\infty}^{\infty} x^2 f(x)dx, untuk\ X\ kontinu \end{cases} \qquad (2.6)$$

(Walpole dan Myers, 1995).

Variabel acak yang mampu menjalanai bilangan bulat adalah variabel acak diskrit, sedangkan variabel acak yang mampu menjalani bilangan *real* adalah variabel acak kontinu.

$\sigma_X =$ Standar Deviasi = $\sqrt{Var(X)}$.

Maka, untuk X variabel random binomial dengan fungsi massa peluang

$$f(X) = b(X; n, p) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x}, & x = 0,1,2,\dots,n \\ 0 & lainnya. \end{cases}$$

$$Var(X) = E[(X - E(X))^2] = E(X^2) - [E(X)]^2 \qquad (2.7)$$

Dari persamaan (2.4) diketahui bahwa

$$E(X) = np$$

$$E(X^2) = \sum_{x=0}^{n} x^2 \binom{n}{x} p^x (1-p)^{n-x}$$

$$= \sum_{x=0}^{n} [x(x-1) - x] \binom{n}{x} p^x (1-p)^{n-x}$$

$$= \sum_{x=0}^{n} x(x-1) \binom{n}{x} p^x (1-p)^{n-x} - \sum_{x=0}^{n} x \binom{n}{x} p^x (1-p)^{n-x}$$

(I) (II)

Dimana, persamaan (II) merupakan persamaan (2.3) yang telah diturunkan menjadi persamaan (2.4)

Persamaan (II)

$$\sum_{x=0}^{n} x \binom{n}{x} p^x (1-p)^{n-x} = np$$

Persamaan (I)

$$\sum_{x=0}^{n} x(x-1) \binom{n}{x} p^x (1-p)^{n-x} = \sum_{x=2}^{n} x(x-1) \binom{n}{x} p^x (1-p)^{n-x}$$

$$= \sum_{x=2}^{n} n(n-1) \binom{n-2}{x-2} p^x (1-p)^{n-x}$$

$$= \sum_{y=0}^{n-2} n(n-1)\binom{n-2}{y} p^{y-2}(1-p)^{n-y-2}$$

$$= n(n-1)p^2 \sum_{y=0}^{n-2} \binom{n-2}{y} p^{y-2}(1-p)^{n-2-y}$$

$$= n(n-1)p^2$$

$$= n^2p^2 - np^2 \tag{2.8}$$

Sehingga berdasarkan persamaan (2.7), maka

$$Var(X) = E[(X - E(X))^2] = E(X^2) - [E(X)]^2$$

$$Var(X) = n^2p^2 - np^2 + np - n^2p^2 = np(1-p) \tag{2.9}$$

(Walpole dan Myers, 1995).

Variabel X merupakan variabel random Binomial dapat ditulis $X \sim Bin(n, p)$.

2.1.6 Distribusi Normal

Distribusi Normal merupakan distribusi teoritis dari variabel random yang kontinu. Distribusi normal merupakan distribusi yang simetris, berbentuk genta dan kontinu serta memiliki fungsi frekuensi

$$f(X) = \frac{1}{\sqrt{2\pi}} exp\left(-\frac{1}{2}\left(\frac{X-\mu_X}{\sigma_x}\right)^2\right) \tag{2.10}$$

(Dajan, 1984).

Fungsi $f(X)$ juga dinamakan fungsi kepekatan normal (normal *density function*). Distribusi normal tergantung pada dua parameter, yaitu rata-rata μ_X dan variansi σ_x^2. karena distribusinya kontinu, cara menghitung probabilitasnya dilakukan dengan jalan menentukan luas di bawah kurvanya. Karena fungsi frekuensi normal tidak memiliki integral yang sederhana, sehingga probabilitas umumnya dihitung dengan menggunakan distribusi normal standar, dimana dapat dicari dengan jalan mengubah variabel random X yang normal ke dalam variabel random Z yang standar dan dirumuskan sebagai

$$Z = \frac{X - \mu_X}{\sigma_x} \tag{2.11}$$

(Dajan, 1984).

Jika Z merupakan variabel random yang kemungkinan harga-harganya menyatakan bilangan-bilangan riil; antara $-\infty$ dan $+\infty$, maka Z dinamakan variabel normal standar jika dan hanya jika probabilitas interval dari a dan b menyatakan luas dari a ke b antara sumbu Z dan kurva normalnya (tabel lengkap Z dapat dilihat pada lampiran 2) dan persamaannya yaitu

$$f(Z) = \frac{1}{\sqrt{2\pi}} exp\left(-\frac{1}{2}Z^2\right) \tag{2.12}$$

(Dajan, 1984).

Persamaan (2.12) dinamakan fungsi kepadatan normal standar (*standard normal density fungtion*). Variabel random Z akan memiliki rata-rata $\mu_X = 0$ dan $\sigma_x^2 = 1$.

2.1.7 Theorema Limit Pusat (*The Central Limit Theorem*)

Theorema limit pusat (*Central Limit Theorem*) menyatakan bahwa jika sampel acak dipilih dari populasi dengan rata-rata μ_x dan σ_x^2 serta jika ukuran sampel n bertambah makin besar, maka rata-rata sampel akan memiliki distribusi pemilihan sampel yang mendekati distribusi normal dengan rata-rata $\mu_{\bar{x}} = \mu_x$ dan standar deviasi $\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}$. Jika populasi terbatas, maka rata-rata sampel akan memiliki distribusi pemilihan sampel dengan rata-rata $\mu_{\bar{x}} = \mu_x$ dan standar deviasi $\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}\sqrt{\frac{(N-n)}{(N-1)}}$. Andaikan populasi tersebut normal maka distribusi pemilihan sampelnya akan normal (Dajan, 1984).

2.1.8 Hubungan antara Distribusi Normal dan Distribusi Binomial

Jika ukuran sampel *n* besar sekali, distribusi binomial dapat disesuaikan sedemikian sehingga dapat didekati dengan distribusi normal standar. Variabel random X atau jumlah sukses dalam n percobaan binomial merupakan penjumlahan dari variabel random sejumlah n dimana tiap variabel random dimaksudkan bagi stiap percobaan binomial dan tiap percobaan menghasilkan nilai 0 atau 1. Dalam keadaan yang biasa, jumlah dari beberapa variabel random selalu mendekati dengan distribusi normal, sehingga distribusi jumlah variabel random dapat didekati dengan distribusi normal bila n makin menjadi besar.

Batas distribusi binomial dapat dipahami secara berangsur-angsur dengan memperhatikan 3 hal pokok berikut (Dajan, 1984).

1. Distribusi binomial merupakan distribusi yang diskrit sedangkan distribusi normal merupakan distribusi yang kontinu, sehingga probabilitas yang dinyatakan dengan ordinat-binomial perlu diganti dengan luas binomial karena luas selalu dipakai untuk menyatakan probabilitas dalam distribusi kontinu.

2. Skala X perlu diganti dengan skala Z. Berdasarkan persamaan (2.4) dinyatakan bahwa untuk distribusi binomial maka $E(X) = np$ dan dari persamaan (2.9) diketahui bahwa $Var(X) = np(1-p)$. Sehingga variabel Z menjadi

$$Z = \frac{X - \mu_X}{\sigma_X} = \frac{X - np}{\sqrt{np(1-p)}} \tag{2.13}$$

Variabel Z di atas mempunyai $\mu_Z = 0$ dan $\sigma_Z^2 = 1$.

3. Pendekatan secara normal terhadap probabilitas binomial dapat dilakukan dengan menghitung luas yang terdapat di bawah kurva normal. Jumlah Probabilitas yang terdapat diantara kurva dan sumbu X adalah sama dengan 1.

2.1.9 Perkiraan Proporsi

Proporsi atau presentase menunjukkan suatu karakteristik atau ciri eksperimen binomial, suatu observasi termasuk atau tidak termasuk dalam kategori tertentu yaitu kategori yang menjadi perhatian. Dalam hal ini dikelompokkan menjadi dua kategori, yaitu memilih atau tidak memilih dalam pelaksanaan pemilu. Notasi yang akan digunakan: *P* menyatakan proporsi populasi dan *p* menyatakan taksiran *P* berdasarkan sampel acak. Proporsi dapat dipandang sebagai hal khusus dari mean dengan tiap variabel random X, akan berharga 0 dan 1. Misalkan $X_i,\ i = 1, \dots, N$ menyatakan nilai populasi, maka proporsi *P* didefinisikan sebagai

$$P = \frac{1}{N}\sum_{i=1}^{N} X_i \tag{2.14}$$

(Cochran, 1977).

Jadi definisi proporsi P sama dengan definisi mean $\bar{X} = \frac{1}{N}\sum_{i=1}^{N} X_i$. Hal ini mengakibatkan bahwa seluruh rumus untuk $\bar{X}$ juga berlaku pada P.

Misalkan $X_1, X_2, X_3, X_4, X_5, \ldots, X_n$ adalah sampel acak yang diambil dari populasi $X_1, X_2, X_3, X_4, X_5, \ldots, X_N$. Maka, proporsi sampel p didefinisikan sebagai

$$p = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{x} \tag{2.15}$$

(Cochran, 1977).

Misalkan $X_1, X_2, X_3, X_4, X_5, \ldots, X_N$ populasi berukuran N, dari populasi tersebut diambil sampel acak $X_1, X_2, X_3, X_4, X_5, \ldots, X_n$. Variansi S^2 adalah

$$S^2 = \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \bar{X})^2 = \frac{1}{N-1}\left[\sum_{i=1}^{N} {X_i}^2 - N\bar{X}^2\right]$$

Karena variabel random X, akan bernilai 0 dan 1, diperoleh

$$\sum_{i=1}^{N} {X_i}^2 = \sum_{i=1}^{N} X_i = NP = A \; dan \; \bar{X}^2 = P^2 \, .$$

Jadi,

$$S^2 = \frac{1}{N-1}[NP - NP^2]$$

$$= \frac{N}{N-1}PQ, \quad dengan \; Q = 1 - P \tag{2.16}$$

(Cochran, 1977).

Variansi sampel s^2 menurut definisi adalah

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{x})^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n} {X_i}^2 - n\bar{x}^2\right]$$

Karena variabel random X_i akan bernilai 0 atau 1, maka

$$\sum_{i=1}^{n} {X_i}^2 = \sum_{i=1}^{n} X_i = np = a \; dan \; \bar{x}^2 = p^2 \,.$$

Jadi,

$$s^2 = \frac{1}{n-1}[np - np^2]$$

$$= \frac{n}{n-1} pq \qquad (2.17)$$

dengan $q = 1 - p$

Variansi untuk rata-rata populasi

$V(\bar{X}) = v(\bar{x})$

$\sigma^2 = E(\bar{x} - \bar{X})^2$

$$= \frac{1}{N}\sum_{i=1}^{N}(X_i - \bar{X})^2 \qquad (2.18)$$

$$S^2 \quad = \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \bar{X})^2$$

$N\sigma^2 = (N-1)S^2$

$$\sigma^2 \quad = \frac{N-1}{N} S^2$$

$v(\bar{x}) = \sigma_{\bar{x}}^2 = E[\bar{x} - E(\bar{x})]^2$

$$= E(\bar{x} - \bar{X})^2$$

$$= E\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})\right)^2$$

$$= \frac{1}{n^2} E\left(\sum_{i=1}^{n}(X_i - \bar{X})\right)^2$$

$$= \frac{1}{n^2} E\left[\sum_{i=1}^{n}(X_i - \bar{X})^2 + \sum_{i \neq j}^{n}(X_i - \bar{X})(X_j - \bar{X})\right]$$

$$= \frac{1}{n^2}\left[\sum_{i=1}^{n} E(X_i - \bar{X})^2 + \sum_{i \neq j}^{n} E(X_i - \bar{X})(X_j - \bar{X})\right]$$

$$= \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2 + \frac{1}{n^2}\sum_{i \neq j}^{n} E(X_i - \bar{X})(X_j - \bar{X})$$

Untuk pengambilan sampel tanpa pengembalian,

$$\sigma_{\bar{x}}^2 = \frac{1}{n^2} n\sigma^2 + \frac{1}{n^2}\sum_{i \neq j}^{n}\frac{1}{N}\frac{1}{N-1}\sum_{i \neq j}^{N}(X_i - \bar{X})(X_j - \bar{X})$$

Ingat

$$\left[\sum_{i=1}^{N}(X_i - \bar{X})\right]^2 = \sum_{i=1}^{N}(X_i - \bar{X})^2 + \sum_{i \neq j}^{N}(X_i - \bar{X})(X_j - \bar{X})$$

$$\sum_{i \neq j}^{N}(X_i - \bar{X})\left(X_j - \bar{X}\right) = -\sum_{i=1}^{N}(X_i - \bar{X})^2$$

sehingga

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} - \frac{1}{n^2}\frac{1}{N}\frac{1}{N-1}\sum_{i \neq j}^{n} N\sigma^2$$

$$= \frac{\sigma^2}{n} - \frac{1}{n^2}\frac{1}{N-1} n(n-1)\sigma^2$$

$$= \frac{\sigma^2}{n}\left[1 - \frac{n-1}{N-1}\right] = \frac{\sigma^2}{n}\frac{N-n}{N-1}$$

$$= \frac{N-1}{N}\frac{S^2}{n}\frac{N-n}{N-1} = \frac{S^2}{n}\frac{N-n}{N}$$

(2.20)

Karena variabel random X_i akan bernilai 0 atau 1, diperoleh dari persamaan (2.14) dan (2.15)

$$P = \frac{1}{N}\sum_{i=1}^{N} X_i = \bar{X}$$

dan

$$p = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}$$

maka untuk pengambilan sampel tanpa pengembalian, berlaku persamaan (2.20). Dengan memasukkan nilai S^2 dari persamaan (2.21), maka diperoleh variansi untuk p.

$$V(p) = V(\bar{x})$$

$$= \frac{S^2}{n}\frac{N-n}{N} = \frac{NPQ}{(N-1)n}\frac{N-n}{N} = \frac{PQ}{n}\frac{N-n}{N-1} \qquad (2.21)$$

Dari persamaan (2.20), umunya S^2 tidak diketahui. Jadi perlu ditentukan taksiran untuk S^2 dengan mengganti S^2, diperoleh taksiran $V(p)$, notasi $\hat{V}(p)$. Taksiran S^2 adalah

$$S^2 = s^2 = \frac{n}{n-1}pq. \qquad (2.22)$$

Jadi untuk pengambilan sampel tanpa pengembalian, menjadi

$$\hat{V}(p) = \frac{pq}{n-1}\frac{N-n}{N}. \qquad (2.23)$$

Dari persamaan (2.21) ataupun taksiran persamaan (2.23) dapat diperoleh konfidensi interval untuk p. $(1-\alpha)100\%$ konfidensi interval untuk p adalah

$$p - Z\sqrt{\hat{V}(p)} < p < p + Z\sqrt{\hat{V}(p)}$$

(2.24)

$$p - Z\sqrt{\frac{pq}{n-1}\frac{N-n}{N}} < p < p + Z\sqrt{\frac{pq}{n-1}\frac{N-n}{N}}$$

dengan $Z = Z_{\alpha/2}$ = koefisien reliabilitas atau nilai tabel normal standar.

Untuk menaksir parameter proporsi dengam *margin of error* dan koefisien reliabelitas tertentu, maka perlu ditentukan ukuran sampel n yang akan diambil rumus dasar antara *margin of error*, koefisien reliabelitas dan standar error yaitu

$$E = Zse_p \text{ atau } E^2 = (Zse_p)^2 \; dengan \; Z = Z_{\alpha/2} \tag{2.25}$$

Standar error untuk proporsi (se_p), diberikan oleh persamaan (2.23), maka persamaan (2.25) untuk *sampling* tanpa pengembalian menjadi

$$E^2 = Z^2\frac{pq}{n}\frac{N-n}{N}$$

$$NnE^2 = Z^2pqN - Z^2pqn$$

$$NnE^2 + Z^2pqn = Z^2pqN$$

$$n(NE^2 + Z^2pq) = Z^2pqN$$

$$n = \frac{NZ^2pq}{NE^2 + Z^2pq} \tag{2.26}$$

(Cochran, 1977)

Jadi, ukuran sampel yang dibutuhkan untuk suatu *margin of error* dan koefisien reliabilitas tertentu adalah

$$n = \frac{NZ^2pq}{NE^2 + Z^2pq}$$

Dalam *quick count*

$n =$ ukuran atau jumlah sampel pemilih dalam pemilu

$Z =$ Koefisien reliabilitas atau nilai variabel normal standar

E = Tingkat kesalahan yang ditoleransi (*margin of error*)

p = proporsi yang memilih dalam pemilu

q = proporsi yang tidak memilih dalam pemilu yaitu $(1-p)$

N = Ukuran atau jumlah populasi pemilih dalam pemilu.

2.1.10 Interval Keyakinan untuk Parameter Statistika

Sebuah parameter populasi mempunyai nilai dengan batas toleransi yang dinyatakan dalam interval keyakinan. Interval keyakinan untuk parameter θ, ditulis dalam bentuk interval $L \leq \theta \leq U$, dimana L adalah batas bawah dan U adalah batas atas dari interval. Panjang interval keyakinan tergantung dari tingkat nyata (*significant*) yang dinyatakan dalam $(1-\alpha)100\%$, yang nilainya ditentukan sesuai dengan keperluan analisis statistika (Bain dan Engelhardt, 1992).

Dalam membuat interval keyakinan untuk parameter θ, batas bawah L dan batas atas U ditentukan dengan membuat

$$P[L \leq \theta \leq U] = 1-\alpha \qquad (2.27)$$

(Bain dan Engelhardt, 1992).

Persamaan (2.27) dibaca sebagai probabilitas parameter θ berada pada daerah interval dengan batas bawah L dan batas atas U sama dengan $1-\alpha$. Sehingga, sebuah interval

$$L \leq \theta \leq U \qquad (2.28)$$

merupakan interval keyakinan $(1-\alpha)100\%$ untuk parameter θ. Batas bawah L dan batas atas U masing-masing disebut sebagai batas keyakinan atas dan batas keyakinan bawah, dan $1-\alpha$ disebut koefisien keyakinan. Interprestasi dari interval keyakinan (2.28) menyatakan bahwa, dari sejumlah sampel randm yang diambil, sebesar $(1-\alpha)100\%$ berada pada interval keyakinan sebagai perkiraan nilai θ yang sebenarnya. Panjang interval untuk suatu parameter populasi adalah suatu ukuran yang menentukan kualitas informasi parameter tersebut yang diperoleh dari sampel.

Interval keyakinan $L \leq \theta \leq U$ merupakan interval keyakinan dua arah dengan batas atas dan batas bawah untuk parameter θ. Disamping interval dua arah, dapat pula ditentukan interval keyakinan satu arah, yaitu hanya menggunakan satu batas interval sebagai batas atas atau batas bawah saja. Sebuah interval keyakinan satu arah dengan batas atas $(1-\alpha)100\%$ untuk θ diberikan dengan interval

$$\theta \leq U \tag{2.29}$$

dimana batas atas U dipilih sehingga $P[\theta \leq U] = 1-\alpha$. Notasi P menunjukkan probabilitas. Sehingga jika diambil $\alpha = 0{,}05$ atau 5%, maka $P[\theta \leq U] = 1 - 0{,}05 = 0{,}95$. Sedangkan untuk interval keyakinan satu arah dengan batas bawah $(1-\alpha)100\%$ utnuk θ diberikan dengan interval

$$L \leq \theta \tag{2.30}$$

dimana batas bawah L dipilih sehingga $P[\theta \leq L] = 1-\alpha$.

2.1.11 Interval Konfidensi Bayes Prior Beta

Inferensi Bayes adalah salah satu cabang ilmu matematika yang mempelajari probabilitas dengan syarat adanya prior dan posterior. Prior adalah suatu probabilitas subjektif yang diyakini akan terjadi, sedangkan posterior adalah suatu probabilitas untuk suatu kejadian dengan syarat kejadian yang lain telah terjadi. Dalam teori probabilitas dan statistik, distribusi Beta adalah distribusi probabilitas kontinu dalam interval $(0,1)$ dengan dua parameter yang positif dan biasanya dinotasikan a dan b. Dalam hal ini distribusi Beta digunakan untuk menjelaskan distribusi dari sebuah nilai probabilitas yang tidak diketahui sebagai distribusi prior pada sebuah parameter probabilitas sukses dalam distribusi Binomial (Bolstad, 2007). Dalam hal ini dianggap bahwa probaibilitas sukses θ dapat menjalani setiap nilai *real* antara 0 dan 1, sehingga distribusi prior tidak diskrit tidak realistis (Soejoeti dan Soebanar, 1988).

Dalam statistik Bayes distribusi Beta dapat dilihat sebagai probaibilitas parameter θ pada distribusi Binomial setelah observasi $a-1$ sukses (dengan probabilitas θ sebagai probabilitas sukses) dan $b-1$ gagal (dengan probabilitas $1-\theta$ gagal) (Bolstad, 2007).

Beta (a, b) sebagai prior memiliki densitas

$$f(\theta: a, b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\theta^{a-1}(1-\theta)^{b-1} \quad untuk\ 0 \leq \theta \leq 1 \qquad (2.31)$$

(Bolstad, 2007).

Dalam distribusi Beta, kuantitas yang tidak diketahui adalah θ dimana merupakan probabilitas sukses dalam distribusi Binomial, sehingga yang membatasi nilai probabilitas ini haruslah dari 0 dampai dengan 1. Maka cukup beralasan untuk menganggap bahwa θ dapat menjalani banyak tak berhingga nilai-nilai *real* dari 0 sampai dengan 1 dan menggunakan distribusi kontinu (seperti distribusi Beta) sebagai distribusi prior (Soejoeti dan Soebanar, 1988).

a. Prosedur Memilih Prior

Teorema Bayes memberikan metode untuk memilih keyakinan terhadap suatu parameter dari sebuah distribusi jika data diperoleh. Karena dalam kasus ini ditetapkan bahwa Beta (a, b) sebagai prior, maka untuk menggunakan teorema ini, harus dipunyai Beta (a, b) yang merepresentasikan keyakinan terhadap parameter tersebut, sehingga ada beberapa pertimbangan dalam menentukan parameter a dan b pada Beta (Bolstad, 2007).

Hal yang diperhatikan dalam menentukan Beta (a, b), yaitu: *Menghitung persamaan ukuran sampel dari prior* (n_{eq}). Diketahui bahwa proporsi Binomial $\theta = \frac{x}{n}$, maka diperoleh mean proporsi Binomial adalah:

$$E(\theta) = E\left(\frac{x}{n}\right)$$

$$= \frac{1}{n}E(x)$$

$$= \frac{1}{n}n\theta$$

$$= \theta$$

dan variansi proporsi Binomial adalah

$$Var(\theta) = Var\left(\frac{x}{n}\right)$$

$$= \frac{1}{n^2}Var(x)$$

$$= \frac{1}{n^2}n\theta(1-\theta)$$

$$= \frac{\theta(1-\theta)}{n}$$

(Bolstad, 2007).

Teorema. (Berger, 1990)

Mean dan variansi dari distribusi Beta dengan parameter a dan b masing-masing adalah

$$\mu = \frac{a}{a+b} \text{ dan } \sigma^2 = \frac{ab}{(a+b+1)(a+b)^2}$$

Bukti

Menghitung momen dari distribusi Beta bisa dilakukan dengan metode sebagai berikut

$$EX^n = \frac{1}{B(a,b)} \int_0^1 x^n x^{a-1} (1-x)^{b-1} dx$$

$$= \frac{1}{B(a,b)} \int_0^1 x^{(a+n)-1} (1-x)^{b-1} dx \qquad (2.32)$$

maka dapat diperoleh juga persamaan

$$EX^n = \frac{B(a+n,b)}{B(a,b)} = \frac{\Gamma(a+n)\Gamma(a+b)}{\Gamma(a+b+n)\Gamma(a)}. \qquad (2.33)$$

Ingat, Fungsi Gamma didefinisikan oleh

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

untuk $n > 0$, n pecahan negatif n bukan bilangan bulat negatif. Dan Suatu variabel acak dikatakan memiliki distribusi Beta dengan parameter a dan b, jika fungsi kepadatannya adalah

$$f(x) = \begin{cases} \frac{1}{B(a,b)} x^{a-1} (1-x)^{b-1} & 0 < x < 1 \\ 0 & \text{yang lain} \end{cases}$$

dimana $B(a,b)$ merupakan fungsi Beta yang didefinisikan sebagai

$$B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}dx \qquad a > 0, b > 0.$$

Fungsi Beta dihubungkan dengan fungsi Gamma oleh

$$B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}.$$

Sehingga distribusi Beta juga dapat didefinisikan oleh fungsi kepadatan

$$f(x) = \begin{cases} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1} & 0 < x < 1 \\ 0 & \text{yang lainnya.} \end{cases}$$

Berdasarkan persamaan (2.32) dan persamaan (2.33), maka untuk memperoleh mean $(E(X))$ dan $var(X) = E(X^2) - [E(X)]^2$ adalah dengan mensubstitusikan $n = 1$ dan $n = 2$ ke persamaan (2.33), maka

$$Mean(X) = EX^1 = \frac{\Gamma(a+1)\Gamma(a+b)}{\Gamma(a+b+1)\Gamma(a)}$$

$$= \frac{a\Gamma(a)\Gamma(a+b)}{(a+b)\Gamma(a+b)\Gamma(a)}$$

$$= \frac{a}{(a+b)} *$$

dan

$$Var(X) = E(X^2) - [E(X)]^2.$$

Karena

$$EX^2 = \frac{\Gamma(a+2)\Gamma(a+b)}{\Gamma(a+b+2)\Gamma(a)}$$

$$= \frac{(a+1)a\Gamma(a)\Gamma(a+b)}{(a+b+1)(a+b)\Gamma(a+b)\Gamma(a)}$$

$$= \frac{(a+1)a}{(a+b+1)(a+b)},$$

maka

$$Var(X) = \frac{(a+1)a}{(a+b+1)(a+b)} - \left(\frac{a}{a+b}\right)^2$$

$$= \frac{(a+1)a}{(a+b+1)(a+b)} - \left(\frac{a^2}{(a+b)^2}\right)$$

$$= \frac{(a+b)(a^2+a) - a^2(a+b+1)}{(a+b)^2(a+b+1)}$$

$$= \frac{ab}{(a+b)^2(a+b+1)} \ *$$

Karena Beta (a, b) merupakan prior dengan mean prior adalah $\frac{a}{a+b}$ dan variansi prior adalah $\frac{ab}{(a+b)^2(a+b+1)}$, maka dengan menyamakan variansi proporsi Binomial dengan varansi prior diperoleh

$$\frac{\theta(1-\theta)}{n_{eq}} = \frac{ab}{(a+b)^2(a+b+1)}.$$

Dengan menyamakan mean prior dan mean proporsi maka diperoleh $\theta = \frac{a}{a+b}$ dan $1 - \theta = \frac{b}{a+b}$, sehingga persamaan ukuran sampel n_{eq} diperoleh

$$\frac{\theta(1-\theta)}{n_{eq}} = \frac{\theta(a+b)(1-\theta)(a+b)}{(a+b+1)(a+b)^2}$$

$$\frac{\theta(1-\theta)}{n_{eq}} = \frac{\theta(1-\theta)}{a+b+1}$$

$$n_{eq} = a+b+1$$

Ini berarti bahwa banyaknya informasi terhadap parameter θ dari prior yang dipilih mendekati banyaknya sampel random. Sehingga harus diketahui apakah informasi prior terhadap θ benar-benar sama terhadap informasi θ, salah satu caranya dengan memeriksa ukuran sampel random n_{eq}.

Jika data yang dimiliki cukup, maka efek terhadap prior yang dipilih akan lebih kecil dibandingkan dengan data. Dengan kata lain bahwa distribusi posterior yang diperoleh akan memproleh hasil yang mirip walaupun menggunakan prior yang berbeda (Bolstad, 2007).

Metode dalam memilih parameter prior Beta(a, b) adalah memilih Prior Konjugat dengan mencocokkan Mean dan Variansi. Distribusi Beta(a,b) adalah prior konjugat untuk distribusi Binomial (n, θ), dimana distribusi Beta(a, b) memiliki beberapa bentuk berdasarkan parameter a dan b yang dipilih, sehingga parameter prior yang dipilih seharusnya merepresentasikan dengan penilaian subjektif peneliti itu sendiri. Salah satu metodenya adalah dengan memilih Beta(a, b) yang cocok dengan keyakinan prior berdasarkan mean dan standard deviasi.

Jika $\theta = \frac{x}{n}$ merupakan proporsi Binomial, maka mean dari proporsi Binomial adalah $E(\theta) = E\left(\frac{x}{n}\right) = \theta$, dan mean Beta$(a, b)$ adalah $\frac{a}{a+b}$. Dengan menyamakan persamaan mean Beta(a, b) sebagai mean proporsi Binomial diperoleh

$$\theta = \frac{a}{a+b} \tag{2.34}$$

sehingga

$$\begin{aligned} 1 - \theta &= 1 - \frac{a}{a+b} \\ &= \frac{a+b}{a+b} - \frac{a}{a+b} \\ &= \frac{b}{a+b}. \end{aligned}$$

Diketahui standard deviasi distribusi Beta(a,b) adalah $\sqrt{\frac{ab}{(a+b+1)(a+b)^2}}$, dimana dengan persamaan (2.34) dapat diperoleh persamaan $a=\theta(a+b)$ dan $b=(a-\theta)(a+b)$, jika $\sigma=\sqrt{\frac{\theta(1-\theta)}{n}}$ merupakan standard deviasi dari proporsi Binomial, dengan menyamakan standar deviasi Beta(a,b) sebagai standar deviasi dari proporsi Binomial, maka σ juga dapat dinyatakan sebagai

$$\sigma=\sqrt{\frac{\theta(a+b)(1-\theta)(a+b)}{(a+b+1)(a+b)^2}}$$

$$=\sqrt{\frac{\theta(1-\theta)}{a+b+1}}$$

Sehingga variansi dari proporsi Binomial juga dapat dinyatakan sebagai

$$\sigma^2=\frac{\theta(1-\theta)}{a+b+1} \tag{2.35}$$

Dengan persamaan (2.34) diperoleh

$$\theta(a+b)=a$$

$$\theta a+\theta b-a=0$$

$$(\theta-1)a+\theta b=0 \tag{2.36}$$

Karena diketahui bahwa θ merupakan proporsi Binomial dimana $\theta=\frac{x}{n}$, maka persamaan (2.36) menjadi

$$\left(\frac{x}{n}-1\right)a+\frac{x}{n}b=0$$

$$\left(\frac{x-n}{n}\right)a+\frac{x}{n}b=0$$

$$(x-n)a+xb=0 \tag{2.37}$$

dan dengan persamaan (2.35) diperoleh

$$(a+b+1)\sigma^2 = \theta(1-\theta)$$

$$\sigma^2 a + \sigma^2 b + \sigma^2 = \theta(1-\theta). \tag{2.38}$$

Karena variansi proporsi Binomial $\sigma^2 = \frac{\theta(1-\theta)}{n}$, maka persamaan (2.38) adalah

$$\frac{\theta(1-\theta)}{n}a + \frac{\theta(1-\theta)}{n}b + \frac{\theta(1-\theta)}{n} = \theta(1-\theta) \tag{2.39}$$

Jika ruas kanan dan ruas kiri pada persamaan (2.39) dikalikan dengan $\frac{1}{\theta(1-\theta)}$, maka

$$\frac{1}{n}a + \frac{1}{n}b + \frac{1}{n} = 1$$

$$a + b = n - 1 \tag{2.40}$$

sehingga jika diketahui x dan n, maka dengan metode eliminasi persamaan (2.37) dan persamaan (2.40) dapat diselesaikan berdasarkan a dan b, maka

$$(x-n)a + xb = 0 \tag{2.41.a}$$

$$a + b = n - 1 \tag{2.41.b}$$

Persamaan (2.41.b) dikalikan dengan x, maka

$$(x-n)a + xb = 0 \tag{2.41.c}$$

$$xa + xb = x(n-1) \tag{2.41.d}$$

Persamaan (2.41.c) dikurangi dengan (2.41.d), maka a diperoleh

$$\big((x-n)-x\big)a = -\big(x(n-1)\big)$$

$$a = \frac{-(x(n-1))}{-n}$$

$$a = \frac{x(n-1)}{n}. \tag{2.42}$$

Dengan mensubstitusikan persamaan (2.42) ke persamaan (2.41.b), maka dapat diperoleh persamaan b sebagai berikut

$$\frac{x(n-1)}{n} + b = n - 1$$

$$b = (n-1) - \left(\frac{x(n-1)}{n}\right)$$

$$b = \frac{n(n-1) - x(n-1)}{n}$$

$$b = \frac{(n-x)(n-1)}{n} \qquad (2.43)$$

Sehingga dengan persamaan (2.42) dan persamaan (2.43) diperoleh parameter Beta(a, b) yang akan digunakan sebagai prior (Bolstad, 2007).

b. Distribusi Posterior dari Distribusi Binomial

Dalam teorema Bayes setelah data diambil dan prior telah ditentukan maka distribusi posteriornya dicari dengan mengalikan priornya dengan *likelihood*nya. Dalam hal ini prior independen terhadap *likelihood*nya, sehingga data yang diobservasi harus independen terhadap prior yang telah ditentukan (Bolstad, 2007). (2.44)

$$f(\theta|x) \propto f(\theta).f(x|\theta)$$

Jika $f(\theta|x)$ merupakan distribusi posterior dari distribusi Binomial dengan prior konjugat, maka distribusi posterior marginal untuk Proporsi Binomial θ adalah (Bolstad, 2007)

$$f(\theta|x) = \int_0^1 f(\theta)f(x|\theta)d\theta \qquad (2.45)$$

Untuk mendapatkan distribusi posterior, maka persamaan (2.44) dibagi dengan beberapa k (konstanta) untuk membuat posterior menjadi distribusi probabilitas, artinya persamaan (2.44) harus dibagi dengan persamaan (2.45) sehingga distribusi posterior dapat dirumuskan sebagai berikut

$$f(\theta|x) = \frac{f(\theta)f(x|\theta)}{\int_0^1 f(\theta)f(x|\theta)d\theta} \tag{2.46}$$

sehingga fungsi integrasi menjadi dependen terhadap prior $f(\theta)$ yang dipilih. Ini adalah teorema Bayes untuk variabel random kontinu (Bolstad, 2007).

Fungsi kepadatan $f(\theta|x)$ dan $f(\theta)$ masing-masing menunjukkan distribusi posterior dan distribusi prior, sedangkan $f(\theta|x)$ menunjukkan fungsi *likelihood*. Istilah-istilah ini mempunyai intrepetasi yang sama untuk variabel random kontinu seperti halnya variabel random diskrit. Distribusi prior dan posterior harus benar-benar merupakan fungsi densitas, yakni posterior harus bernilai tidak negatif dan jumlah luasan dibawah kurva yang ditentukan dengan pengintegralan fungsi kepadatan itu meliputi seluruh domainnya serta harus sama dengan satu. Sehingga persamaan (2.44) membuat distribusi posterior benar-benar merupakan distribusi probabilitas, dengan fungsi *likelihood* adalah fungsi θ dengan x diketahui (sama dengan nilai observasi x) (Soejoeti dan Soebanar, 1988).

Harus diperhatikan bahwa pada persamaan (2.44) dianggap bahwa θ adalah variabel random kontinu. Jika variabel random yang menjadi perhatian adalah distribusi Binomial dan informasi sampel terdiri dari banyaknya “sukses” dalam n percobaan tertentu maka model probabilitas tersebut adalah diskrit, sedangkan distribusi prior θ dapat berupa diskrit atau kontinu. Istilah "teorema Bayes untuk model probabilitas diskrit" dan "teorema Bayes untuk model probabilitas kontinu" menunjukkan kepada bentuk distribusi prior dan posterior (yakni menunjukkan apakah variabel random θ dianggap diskrit atau kontinu) (Soejoeti dan Soebanar, 1988).

2.1.12 Estimator Bayes dari Distribusi Binomial dengan Prior Beta

Jika $X \sim Binomial\ (n, \theta)$ dan densitas prior $\theta \sim Beta(a, b)$, maka fungsi densitas posterior dapat dinyatakan sebagai fungsi bersyarat dari θ dengan x diketahui, berdasarkan definisi : "Jika X_1 dan X_2 merupakan variabel random diskrit atau kontinu dengan fungsi densitas peluang bersama $f(x_1, x_2)$, maka fungsi densitas peluang bersyarat dari X_2 jika diketahui $X_1 = x_1$ didefinisikan dengan:

$$f(x_2|x_1) = \frac{f(x_1, x_2)}{f_1(x_1)}$$

untuk nilai x_1 sedemikian hingga $f_1(x_1) > 0$, dari nol untuk lainnya. Sedangkan fungsi densitas peluang bersyarat dari X_1 jika diketahui $X_2 = x_2$ didefinisikan dengan:

$$f(x_1|x_2) = \frac{f(x_1, x_2)}{f_2(x_2)}$$

untuk nilai x_2 sedemikian hingga $f_2(x_2) > 0$, dari nol untuk lainnya".

Sehingga fungsi bersyarat dari θ dengan x diketahui dapat ditulis dengan

$$f(\theta|x) = \frac{f(\theta, x)}{f(x)} \tag{2.47}$$

Karena $f(\theta, x)$ dapat dinyatakan sebagai $f(x)f(\theta|x)$ atau $f(\theta)f(x|\theta)$, maka

$$f(\theta)f(x|\theta) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\theta^{a-1}(1-\theta)^{b-1}\binom{n}{x}\theta^{x}(1-\theta)^{n-x}$$

$$= \binom{n}{x}\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\theta^{x+a-1}(1-\theta)^{n-x+b-1} \quad (2.48)$$

dimana

$$\binom{n}{x} = \frac{n!}{(n-x)!\,x!} \quad (2.49)$$

Karena $\Gamma(n) = (n-1)!$, maka persamaan (2.49) dapat ditulis sebagai

$$\binom{n}{x} = \frac{\Gamma(n+1)}{\Gamma(\mathrm{n}-\mathrm{x}+1)\Gamma(\mathrm{x}+1)}. \quad (2.50)$$

sehingga dengan mensubstitusikan persamaan (2.50) ke persamaan (2.48) maka diperoleh

$$f(\theta)f(\mathrm{x}|\theta) = \frac{\Gamma(n+1)}{\Gamma(\mathrm{n}-\mathrm{x}+1)\Gamma(\mathrm{x}+1)}\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\theta^{x+a-1}(1-\theta)^{n-x+b-1}$$

Selanjutnya perhatikan $f(x)$, dimana $f(x)$ merupakan fungsi densitas peluang marginal dari x, sehingga

$$f(x) = \int_0^1 f(\theta, x)d\theta$$

$$= \int_0^1 f(\theta)f(x|\theta)$$

$$= \int_0^1 \frac{\Gamma(n+1)}{\Gamma(\mathrm{n}-\mathrm{x}+1)\Gamma(\mathrm{x}+1)}\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\theta^{x+a-1}(1-\theta)^{n-x+b-1}d\theta$$

$$= \frac{\Gamma(n+1)\Gamma(a+b)}{\Gamma(\mathrm{x}+1)\Gamma(\mathrm{n}-\mathrm{x}+1)\Gamma(a)\Gamma(b)}\frac{\Gamma(x+a)\Gamma(n+b-x)}{\Gamma(a+n+b)}$$

(2.51)

$$* \int_0^1 \frac{\Gamma(a+n+b)}{\Gamma(x+a)\Gamma(n+b-x)} \theta^{x+a-1}(1-\theta)^{n+b-x-1} d\theta$$

Perhatikan

$$\frac{\Gamma(a+n+b)}{\Gamma(x+a)\Gamma(n+b-x)} \theta^{x+a-1}(1-\theta)^{n+b-x-1}$$

merupakan integrasi dari fungsi densitas Beta$(x+a, n+b-x)$. Karena x variabel random kontinu, maka fungsi densitas peluangnya akan memenuhi kondisi bahwa $\int_0^1 f(x)dx = 1$, sehingga

$$\int_0^1 \frac{\Gamma(a+n+b)}{\Gamma(x+a)\Gamma(n+b-x)} \theta^{x+a-1}(1-\theta)^{n+b-x-1} d\theta = 1$$

Oleh karena itu persamaan (2.51) dapat ditulis menjadi

$$f(x) = \frac{\Gamma(n+1)\Gamma(a+b)}{\Gamma(x+1)\Gamma(n-x+1)\Gamma(a)\Gamma(b)} \frac{\Gamma(x+a)\Gamma(n+b-x)}{\Gamma(a+n+b)} \quad (2.52)$$

Maka dengan persamaan (2.47), (2.48), dan (2.52) fungsi densitas posterior dapat ditulis sebagai

$$f(\theta|x) = \frac{f(\theta).f(x|\theta)}{\int_0^1 f(\theta)f(x|\theta)d\theta}$$

$$= \frac{\frac{\Gamma(n+1)}{\Gamma(n-x+1)\Gamma(x+1)} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \theta^{x+a-1}(1-\theta)^{n-x+b-1}}{\frac{\Gamma(n+1)\Gamma(a+b)}{\Gamma(x+1)\Gamma(n-x+1)\Gamma(a)\Gamma(b)} \frac{\Gamma(x+a)\Gamma(n+b-x)}{\Gamma(a+n+b)}}$$

$$= \frac{\Gamma(a+n+b)}{\Gamma(x+a)\Gamma(n+b-x)} \theta^{x+a-1}(1-\theta)^{n+b-x-1} \quad (2.53)$$

Berdasarkan persamaan (2.53), dapat diketahui bahwa posterior berdistribusi Beta $(x+a, n-x+b)$ dengan θ merupakan variabel dan x adalah nilai observasi atau sampel.

Dalam perspektif Bayes, estimasi titik mempunyai pengertian bahwa distribusi posterior akan digambarkan oleh nilai dari sebuah statistik tunggal. Nilai yang paling penting untuk menggambarkan distribusi posterior adalah ukuran lokasi. Oleh karena itu, mean posterior dan median posterior disini akan menjadi kandidat terbaik untuk dijadikan sebagai sebuah estimator. Mengacu pada persamaan (2.45) yang membuktikan bahwa mean posterior merupakan estimator yang optimum untuk θ, maka dalam hal ini mean posterior digunakan sebagai estimator Bayes (Bolstad, 2007), sehingga estimator Bayes untuk parameter θ jika dinyatakan sebagai $\hat{\theta}_B$ adalah

$$\hat{\theta}_B = \frac{x+a}{a+b+n} \tag{2.54}$$

Perhatikan estimator Bayes $\hat{\theta}_B$ yang diperoleh, diketahui bahwa distribusi Beta(a,b) yang digunakan sebagai prior yang mempunyai mean $a/(a+b)$ dan $\hat{\theta} = \frac{x}{n}$ yang merupakan proporsi distribusi Binomial, maka estimator Bayes $\hat{\theta}_B$ akan mengkombinasikan estimator $\hat{\theta}$ dengan informasi prior, hal ini terlihat jika $\hat{\theta}_B$ ditulis sebagai

$$\hat{\theta}_B = \left(\frac{n}{a+bn}\right)\left(\frac{x}{n}\right) + \left(\frac{a+b}{a+b+n}\right)\left(\frac{a}{a+b}\right) \tag{2.55}$$

Sehingga pada persamaan (2.55) terlihat bahwa $\hat{\theta}_B$ adalah kombinasi linear dari mean prior dan mean sampel (Berger, 1990).

a. Nilai Ekspektasi dan Variansi Posterior

Teorema :

Mean dan variansi dari distribusi Beta dengan parameter a dan b masing-masing adalah

$$\mu = \frac{a}{a+b} \text{ dan } \sigma^2 = \frac{ab}{(a+b+1)(a+b)^2}$$

Bukti

Menghitung momen dari distribusi Beta bisa dilakukan dnegan metode sebagai berikut

$$EX^n = \frac{1}{B(a,b)}\int_0^1 x^n x^{a-1}(1-x)^{b-1}dx$$

$$= \frac{1}{B(a,b)}\int_0^1 x^{(a+n)-1}(1-x)^{b-1}dx \qquad (2.56)$$

maka dapat diperoleh juga persamaan

$$EX^n = \frac{B(a+n,b)}{B(a,b)} = \frac{\Gamma(a+n)\Gamma(a+b)}{\Gamma(a+b+n)\Gamma(a)}. \qquad (2.57)$$

Berdasarkan persamaan (2.56) dan persamaan (2.57), maka untuk memperoleh mean ($E(X)$) dan $var(X) = E(X^2) - [E(X)]^2$ adalah dengan mensubstitusikan $n = 1$ dan $n = 2$ ke persamaan (2.57), maka

$$Mean(X) = EX^1 = \frac{\Gamma(a+1)\Gamma(a+b)}{\Gamma(a+b+1)\Gamma(a)}$$

$$= \frac{a\Gamma(a)\Gamma(a+b)}{(a+b)\Gamma(a+b)\Gamma(a)}$$

$$= \frac{a}{(a+b)} *$$

dan

$$Var(X) = E(X^2) - [E(X)]^2.$$

Karena

$$EX^2 = \frac{\Gamma(a+2)\Gamma(a+b)}{\Gamma(a+b+2)\Gamma(a)}$$

$$= \frac{(a+1)a\Gamma(a)\Gamma(a+b)}{(a+b+1)(a+b)\Gamma(a+b)\Gamma(a)}$$

$$= \frac{(a+1)a}{(a+b+1)(a+b)},$$

maka

$$Var(X) = \frac{(a+1)a}{(a+b+1)(a+b)} - \left(\frac{a}{a+b}\right)^2$$

$$= \frac{(a+1)a}{(a+b+1)(a+b)} - \left(\frac{a^2}{(a+b)^2}\right)$$

$$= \frac{(a+b)(a^2+a) - a^2(a+b+1)}{(a+b)^2(a+b+1)}$$

$$= \frac{ab}{(a+b)^2(a+b+1)} \; *$$

Berdasarkan teorema nilai ekspektasi (μ) dan variansi (σ^2) dari distribusi Beta(a, b) adalah

$$\mu = \frac{a}{a+b} \quad \text{dan} \quad \sigma^2 = \frac{ab}{(a+b+1)(a+b)^2}$$

Diketahui distribusi posterior yang diperoleh berdistribusi Beta $(x+a, n-x+b)$, misalkan $a' = x + a \quad dan \quad b' = n - x + b$, maka diperoleh nilai ekspektasi dari distribusi posteriornya adalah

$$E(\theta|x) = \hat{\theta}_B = \frac{a'}{a' + b'}$$

$$= \frac{x+a}{(x+a)+(n-x+b)}$$

$$= \frac{x+a}{a+b+n} \tag{2.58}$$

dan variansi dari distribusi posterior adalah

$$Var(\theta|x) = \frac{a'b'}{(a'+b')^2(a'+b'+1)}$$

$$= \frac{(x+a)(n-x+b)}{\left((x+a)+(n-x+b)\right)^2((x+a)(n-x+b)+1)}$$

$$= \frac{(x+a)(n-x+b)}{(x+a+n-x+b)^2(x+a+n-x+b+1)}$$

$$= \frac{(x+a)(n-x+b)}{(a+n+b)^2(a+b+n+1)} \qquad (2.59)$$

(Bolstad, 2007).

Bukti persamaan 2.58

Nilai ekspektasi posterior dapat diperoleh dengan

$$E(\theta|x) = \int_0^1 \theta f(\theta|x) d\theta$$

$$= \int_0^1 \theta \frac{\Gamma(a+n+b)}{\Gamma(x+a)\Gamma(n+b-x)} \theta^{x+a-1}(1-\theta)^{n+b-x-1} d\theta$$

$$= \frac{\Gamma(a+n+b)}{\Gamma(x+a)\Gamma(n+b-x)} \int_0^1 \theta[\theta^{x+a-1}(1-\theta)^{n+b-x-1}] d\theta$$

$$= \frac{\Gamma(a+n+b)}{\Gamma(x+a)\Gamma(n+b-x)} \int_0^1 \theta^{x+a}(1-\theta)^{n+b-x-1} d\theta$$

$$= \frac{\Gamma(a+n+b)}{\Gamma(x+a)\Gamma(n+b-x)} \frac{\Gamma(x+a+1)\Gamma(n-x+b)}{\Gamma(x+a+1+n-x+b)}$$

$$= \frac{\Gamma(a+n+b)}{\Gamma(x+a)\Gamma(n+b-x)} \frac{\Gamma(x+a+1)\Gamma(n-x+b)}{\Gamma(a+1+n+b)}$$

$$= \frac{\Gamma(a+n+b)(x+a)\Gamma(x+a)}{\Gamma(x+a)\Gamma(a+1+n+b)}$$

$$= \frac{\Gamma(a+n+b)(x+a)}{\Gamma(a+1+n+b)}$$

$$= \frac{\Gamma(a+n+b)(x+a)}{(\mathrm{n}+\mathrm{a}+\mathrm{b})\Gamma(n+a+b)}$$

$$= \frac{x+a}{\mathrm{n}+\mathrm{a}+\mathrm{b}} *$$

Bukti persamaan 2.59

jika

$$E(\theta^2|x) = \int_0^1 \theta^2 f(\theta|x) d\theta$$

$$= \frac{\Gamma(a+n+b)}{\Gamma(x+a)\Gamma(n+b-x)} \int_0^1 \theta^2 [\theta^{x+a}(1-\theta)^{n+b-x-1}] d\theta$$

$$= \frac{\Gamma(a+n+b)}{\Gamma(x+a)\Gamma(n+b-x)} \int_0^1 \theta^{x+a+1}(1-\theta)^{n+b-x-1} d\theta$$

$$= \frac{\Gamma(a+n+b)}{\Gamma(x+a)\Gamma(n+b-x)} \frac{\Gamma(x+a+2)\Gamma(n+b-x)}{\Gamma(x+a+2+n+b-x)}$$

$$= \frac{\Gamma(a+n+b)\Gamma(x+a+2)}{\Gamma(x+a)\Gamma(a+b+n+2)}$$

$$= \frac{\Gamma(a+n+b)(x+a+1)(x+a)\Gamma(x+a)}{\Gamma(x+a)(a+b+n+1)(a+b+n)\Gamma(a+b+n)}$$

$$= \frac{(x+a+1)(x+a)}{(\mathrm{a}+\mathrm{b}+\mathrm{n}+1)(a+b+n)}$$

Dengan teorema : “Jika X adalah suatu variabel random dengan fungsi densitas peluang $f(x)$, maka variansi dari X yang dinotasikan dengan σ^2 adalah

$$\sigma^2 = E(X-\mu)^2 = E(X^2) - \mu^2 = E(X^2) - [E(X)]^2$$

dimana $\mu = E(X)$".

Bukti

$$E(X-\mu)^2 = E(X^2 - 2\mu X + \mu^2)$$
$$= E(X^2) - 2\mu E(X) + \mu^2$$
$$= E(X^2) - 2\mu^2 + \mu^2$$
$$= E(X^2) - \mu^2$$
$$= E(X^2) - [E(X)]^2 *$$

diperoleh variansi posterior sebagai berikut

$$Var(\theta|x) = E(\theta^2|x) - E(\theta|x)^2$$

$$= \frac{(x+a+1)(x+a)}{(\mathrm{a}+\mathrm{b}+\mathrm{n}+1)(a+b+n)} - \left(\frac{x+a}{a+b+n}\right)^2$$

$$= \frac{(x+a+1)(x+a)}{(\mathrm{a}+\mathrm{b}+\mathrm{n}+1)(a+b+n)} - \frac{(x+a)^2}{(a+b+n)^2}$$

$$= \frac{(x+a+1)(x+a)(a+b+n) - (x+a)^2(a+b+n+1)}{(a+b+n+1)(a+b+n)^2}$$

$$= \frac{(x+a)(((x+a+1)(a+b+n)) - ((x+a)(a+b+n+1)))}{(a+b+n+1)(a+b+n)^2}$$

$$= \frac{(x+a)((xn+xa+xb+an+a^2+ab+n+a+b) - (xa+xb+xn+x+a^2+ab+an+a))}{(a+b+n+1)(a+b+n)^2}$$

$$= \frac{(x+a)(n-x+b)}{(\mathrm{a}+\mathrm{b}+\mathrm{n}+1)(a+b+n)^2} *$$

(Bolstad, 2007).

b. Interval Konfidensi Bayes

Dalam inferensi bayes, interval konfidensi merupakan interval probabilitas posterior yang digunakan untuk estimasi interval, sedangkan pada pendekatan klasik interval konfidensi diperoleh dari data sampel (Bolstad, 2007).

Misalkan $X_1, X_2, \ldots, X_n$ variabel random yang diambil dari suatu populasi sembarang yang mempunyai μ dan variansi σ^2, distribusi sampling untuk mean dapat dianggap mendekati normal dengan $\mu_{\bar{x}} = \mu$ dan variansi $\sigma_{\bar{x}}^2 = \sigma^2$. (2.60) dengan Teorema limit pusat diperoleh

$$Z = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \sim N(0.1)$$

Untuk $n \to \infty$ (Bain dan Engelhardt, 1992). Jika dianggap bahwa σ diketahui maka interval konfidensi untuk μ dengan koefisien konfidensi mendekati $1 - \alpha$ adalah

$$P\left(-Z_{\alpha/2} \leq Z \leq Z_{\alpha/2}\right) \sim 1 - \alpha$$

$$P\left(-Z_{\alpha/2} \leq \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \leq Z_{\alpha/2}\right) \sim 1 - \alpha$$

$$P\left(-Z_{\alpha/2} \leq \frac{(\bar{X}_n - \mu)}{\sigma/\sqrt{n}} \leq Z_{\alpha/2}\right) \sim 1 - \alpha$$

$$P\left(-Z_{\alpha/2}\, \sigma/\sqrt{n} \leq \bar{X}_n - \mu \leq Z_{\alpha/2}\, \sigma/\sqrt{n}\right) \sim 1 - \alpha$$

$$P\left(\bar{X}_n - Z_{\alpha/2}\, \sigma/\sqrt{n} \leq \mu \leq \bar{X}_n + Z_{\alpha/2}\, \sigma/\sqrt{n}\right) \sim 1 - \alpha \qquad (2.61)$$

Diketahui distribusi posterior $f(\theta|x)$ berdistribusi Beta $(x + a, n - x + b)$, dengan $\bar{X}_n = m' = \hat{\theta}_B$ dan $\sigma_{\bar{x}}^2 = s'$, maka interval konfidensi untuk mean $\hat{\theta}_B$ dengan kepercayaan mendekati $1 - \alpha$ dari distribusi posterior Beta $(x + a, n - x + b)$ juga dapat diperoleh dengan mengaproksimasi ke disribusi Normal $N[m'; (s')^2]$, sehingga dapat ditulis sebagai

$$[m' - Z_{\frac{\alpha}{2}}\, s', m' + Z_{\frac{\alpha}{2}}\, s'] \qquad (2.62)$$

Dimana $Z_{\frac{\alpha}{2}}$ adalah nilai tabel Normal standar, mean posterior

$$m' = \frac{x+a}{n+a+b} \tag{2.63}$$

Dan variansi posterior

$$(s')^2 = \frac{(x+a)(n-x+b)}{(a+b+n)^2(a+b+n+1)} \tag{2.64}$$

(Bolstad, 2007).

2.1.13 Uji Hipotesis Bayes

Hipotesis statistik adalah suatu anggapan atau pernyataan yang mungkin benar atau tidak, mengenai satu populasi atau lebih (Walpole dan Myers, 1995). Dalam masalah menguji hipotesis, distribusi posterior digunakan untuk menghitung probabilitas H_0 dan H_1 adalah benar. Tapi perlu diperhatikan bahwa $f(\theta|x)$ merupakan sebuah probabilitas untuk sebuah variabel random. Karenanya, probabilitas posterior $P\left(\theta\epsilon\hat{\theta}_B\middle|x\right) = P(H_0\ benar|x)$ dapat dihitung (Berger, 1990).

Pengujian yang dihadapi merupakan uji dua arah dan satu arah untuk hipotesis H_0 dengan tandingan H_1 dengan taraf signifikansi α.

A. $H_0: \theta\ \leq \theta_0\ \ vs\ \ H_1: \theta > \theta_0$

B. $H_0: \theta\ \geq \theta_0\ \ vs\ \ H_1: \theta < \theta_0$

C. $H_0: \theta = \theta_0\ \ vs\ \ H_1: \theta \neq \theta_0$

Pendekatan yang biasa digunakan pada kasus ini adalah dilakukan dengan mengaproksimasi ke distribusi Normal. Dengan statistik uji pada taraf α, maka kriteria uji

$$Z = \frac{\sqrt{n}(\bar{X}_n - \theta_0)}{\sigma}$$

atau

$$Z = \frac{(\bar{X}_n - \theta_0)}{s/\sqrt{n}}$$

Dengan $\bar{X}_n = m' = \hat{\theta}_B$ dan $\sigma_{\bar{x}}^2 = \sigma/n = s'$, maka

$$Z = \frac{(m' - \theta_0)}{s'}$$

Sehingga tolak H_0 (berdasarkan α dan hipotesis)

A. H_0 ditolak apabila $Z > Z_\alpha$

B. H_0 ditolak apabila $Z < -Z_\alpha$

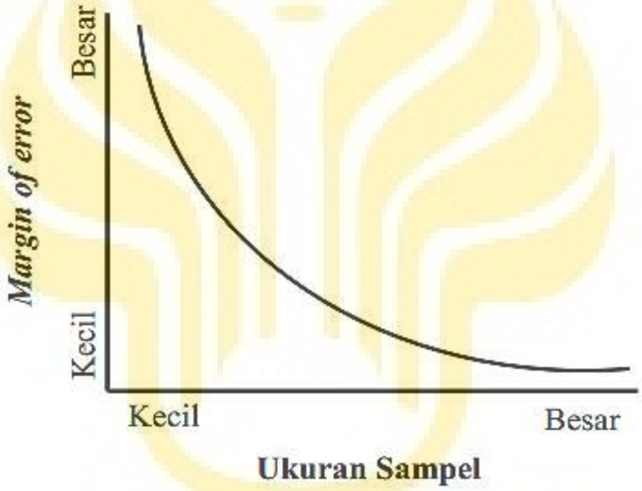
C. H_0 ditolak apabila $Z > Z_{\alpha/2}$ atau $Z < Z_{\alpha/2}$

(Bolstad, 2007).

2.1.14 Tingkat Kesalahan yang Ditoleransi (*Margin of Error*)

Tingkat akurasi diukur diantaranya dari sejauh mana ketepatan sampel dalam menggambarkan populasi. Presisi merupakan pernyataan sejauh mana perbedaan antara nilai statistik dengan nilai parameter. Parameter adalah ciri-ciri yang menjelaskan populasi, sedangkan statistik adalah ciri-ciri yang menjelaskan sampel. Presisi tergantung pada ukuran sampel. Dalam sampel probabilitas, ukuran sampel yang besar akan memberikan presisi yang lebih besar, karena dapat menurunkan kesalahan kesempatan dalam pengacakan (Cochran, 1977).

Hubungan antara ukuran sampel dan tingkat kesalahan dapat digambarkan dalam Gambar 2.1 berikut. Dalam gambar terlihat dalam penelitian yang menggunakan sampel, kurva tidak mungkin menyentuh sumbu X ataupun sumbu Y. Jika kurva menyentuh sumbu X berarti sensus-kesalahan akibat pengambilan sampel adalah 0%. Kurva juga tidak akan menyentuh sumbu Y yang berarti tidak ada sampel-kesalahan 100%. *Margin of error* adalah fungsi terbalik dari ukuran sampel, semakin besar ukuran sampel yang dipakai, nilai *margin of error* semakin kecil (Dajan, 1984).



Gambar 2.1 Hubungan Antara *Margin of Error* dengan Ukuran Sampel

Margin of error pada tingkat kepercayaan $(1-\alpha)100\%$ dapat dihitung dengan mengalikan nilai standar error dengan nilai Z pada tingkat kepercayaan tertentu (Eriyanto, 1990).

$$Margin\ of\ error = Standar\ error\ \times nilai\ Z \tag{2.65}$$

Nilai *standar error* dari *sampling* adalah standar deviasi dari distribusi sampel yang secara teoritik akan terjadi jika diambil sampel dengan ukuran dan populasi sama. Dalam mengestimasi parameter populasi, diasumsikan bahwa sampel akan jauh dalam suatu daerah tertentu pada distribusi sampel. Diasumsikan bahwa *mean* tidak tunggal, tetapi banyak dan menghasilkan hasil yang berbeda bahkan secara ekstrim bisa sangat berbeda dari nilai populasi. Meskipun demikian, prinsip probabilitas menyatakan bahwa mean dari suatu distribusi sampel adalah sama dengan mean populasi, pengulangan, dan pengambilan sampel secara acak akan menghasilkan *mean* yang mengelompok di sekitar nilai mean yang sesungguhnya pada populasi. Dengan kata lain, meskipun hasil mean antar satu sampel dengan sampel lain berbeda, nilai rata-rata dari mean sampel akan sama dengan mean populasi.

Standar deviasi dari populasi tidak diketahui, sehingga digunakan standar deviasi dari sampel, dengan mengasumsikan bahwa standar deviasi populasi sama dengan standar deviasi sampel (Eriyanto, 1999).

Standar error (se_p) untuk proporsi dapat dirumuskan sebagai berikut.

$$se_p = \sqrt{\frac{pq}{n}\frac{N-n}{N}} \qquad (2.66)$$

Dalam *quick count* :

$p =$ proporsi yang memilih dalam pemilu

$q =$ proporsi yang tidak memilih dalam pemilu

$N =$ ukuran populasi pemilih dalam pemilu

$n =$ ukuran sampel pemilih dalam pemilu (Eriyanto, 1990).

2.1.15 Tingkat Kepercayaan

Dalam menentukan ukuran sampel, juga diperhatikan tingkat kepercayaan yang harus diberikan dalam menyimpulkan bahwa jika dilakukan penarikan sampel yang lain dari populasi itu, sampel tersebut harus memberikan hasil yang kira-kira sama dengan pengambilan sampel yang pertama. Diulang berapapun pengambilan sampel, akan memberikan hasil yang sama. Tingkat kepercayaan ini erat hubungannya dengan *margin of error*. *Margin of error* mengacu pada bagaimana keakuratan taksiran yang diinginkan, sedangkan tingkat kepercayaan mengacu kepada bagaimana kepastian yang diinginkan bahwa taksiran itu sendiri akurat.

Tingkat kepercayaan yang sering dipakai adalah 90%, 95%, dan 99%. Disini dapat diyakini bahwa 90% atau 95% bahwa komposisi sampel bisa diulang serta tetap identik jika dipilih sampel lain dari populasi yang sama. Semakin tinggi tingkat kepercayaan yang diinginkan, semakin besar ukuran sampel yang diperlukan. Tingkat kepercayaan dapat memberikan keyakinan bahwa temuan-temuan dalam sampel dapat digeneralisasikan kepada populasi. Jika digunakan tingkat kepercayaan 90% atau 95%, ini berarti jika terdapat 100 sampel, maka perbedaan yang diamati akan muncul dalam 90 atau 95 dari sampel itu (Eriyanto, 1990).

2.1.16 Analisis *Quick Count*

Prediksi *quick count* akan akurat jika mengacu pada metodologi statistik dan penarikan sampel yang ketat serta diimplementasikan secara konsisten di lapangan (*Estok et al*, 2002). Kekuatan *quick count* juga sangat tergantung pada identifikasi terhadap berbagai faktor yang berdampak pada distribusi suara dalam populasi memilih. Apabila pemilu berjalan lancar tanpa kecurangan, akurasi *quick count* dapat disandarkan pada perbandingannya dengan hasil resmi KPU. Tetapi apabila berjalan penuh kecurangan, maka hasil *quick count* dapat dikatakan kredibel meskipun hasilnya berbeda dengan hasil resmi KPU (Ujiyati, 2004).

Keberhasilan *quick count* dinilai dari dua hal (LSI, 2006), yaitu

1. Akurasi

Perhitungan quick count dikatakan memiliki tingkat akurasi yang tinggi apabila hasil perhitungan *quick count* dapat meramalkan pemenang dan urutan komposisi pemenang Pemilu.

2. Presisi

Perhitungan *quick count* dikatakan memiliki tingkat presisi yang tinggi apabila memiliki selisih proporsi yang kecil untuk masing-masing partai atau kandidat calon peserta Pemilu antara hasil perhitungan quick count dengan hasil perhitungan akhir penyelenggara pemilihan umum atau KPU. Suatu *quick count* dapat dikatakan berhasil jika memiliki selisih hasil perhitungan yang lebih kecil dari pada tingkat kesalahan yang ditoleransi (*margin of error*).

Dalam praktiknya, proses perhitungan suara pemilih dalam pemilu KPU dan LSI dapat dibedakan berdasarkan ukuran TPS, petugas atau organisasi penyelenggara, parameter, dan waktu yang dibutuhkan untuk mempublikasikan hasil perolehan suara. Berikut dapat dilihat perbandingan cara perhitungan suara yang dilakukan oleh KPU dan perhitungan oleh LSI untuk mencapai hasil akhir dari proses perhitungan suara sah pemilu.

Tabel 2.1 Perbandingan Proses Perhitungan Suara KPU dengan LSI

Proses Perhitungan Suara	**KPU**	**LSI**
Jumlah TPS	Semua TPS yang ada	TPS yang dipilih secara acak
Petugas	KPU	Relawan LSI
Parameter	Berdasarkan rekapan data yan telah dihitung dari TPS, lalu Kelurahan, Kecamatan, Kabupaten, Provinsi, dan Pusat	Berdasarkan data yang diinput oleh relawan yang dikirim ke server pusat
Waktu	Beberapa hari bahkan beberapa minggu setelah proses perhitungan suara selesai	2-3 jam setelah proses perhitungan suara selesai

Sumber : LSI, 2006

2.1.17 Organisasi *Quick Count*

Untuk dapat melaksanakan *quick count* dengan baik, maka dibutuhkan suatu organisasi yang baik pula. Organisasi yang baik terdiri atas komponen-komponen yang saling terikat dan saling membutuhkan satu sama lainnya, sehingga dibutuhkan kerjasama yang apik antara satu dengan yang lainnya.

Berikut contoh organisasi *quick count* (*Estok et al*, 2002).

1. Media Team

Bertugas menyiapkan sarana dan prasarana dalam proses *quick count*, mulai dari tahap perencanaan, persiapan, pelaksanaan hingga evaluasi/pelaporan kegiatan. Mengembangkan relasi/*networking* dengan media massa, para kandidat, komunitas diplomatik, dan organisasi Internasional. Selain itu juga bertugas untuk mempromosikan *quick count*.

2. Teknikal Team

Bertugas sebagai team teknis dalam pelaksanaan *quick count*. Menyiapkan metode-metode statistik yang akan digunakan, pengumpulan data, serta teknologi yang akan dipakai dalam *quick count*.

3. Administrasi Team

Menyiapkan dan mengatur dokumen-dokumen yang dibutuhkan, serta mengatur *schedull* kerja dan pembagian waktu (*timing*) dalam *quick count*.

4. Relawan/*Surveyor* Team

Sebagai pelaksana lapangan, dalm hal ini relawan bertugas untuk mengambil data di lapangan, berupa hasil perhitungan suara pada masing-masing TPS yang menjadi sampel, serta bertugas menyampaikan data-data tersebut ke pusat data/*database*.

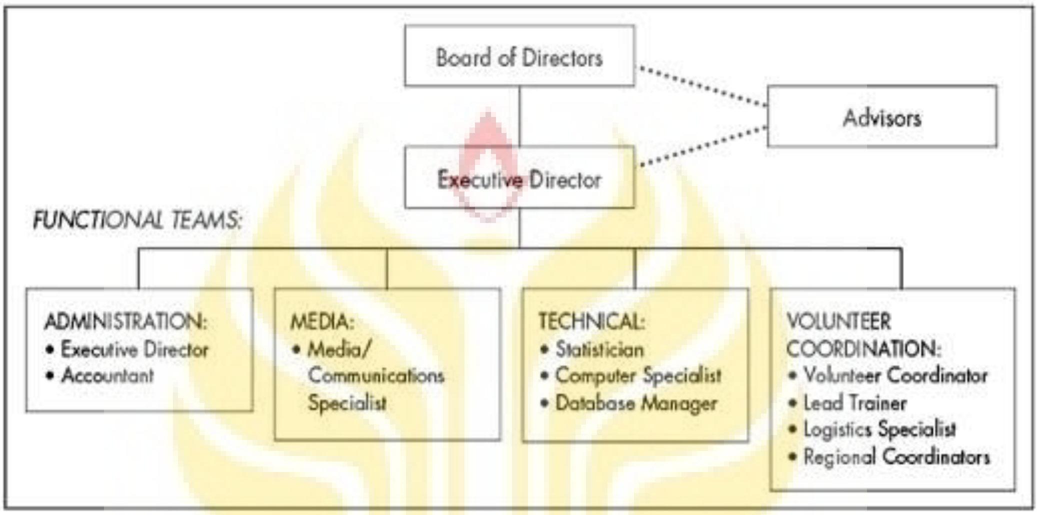
5. Analisis Team

Memanajemen data hasil *quick count*, melakukan analisis terhadap hasil *quick count*, dan menyiapkan publikasi hasil *quick count*.

6. Pimpinan/Direksi

Memanajamen seluruh anggota *team*, melakukan pembagian tugas, menyiapkan laporan hasil *quick count*, dan bertanggungjawab terhadap seluruh kegiatan *quick count*.

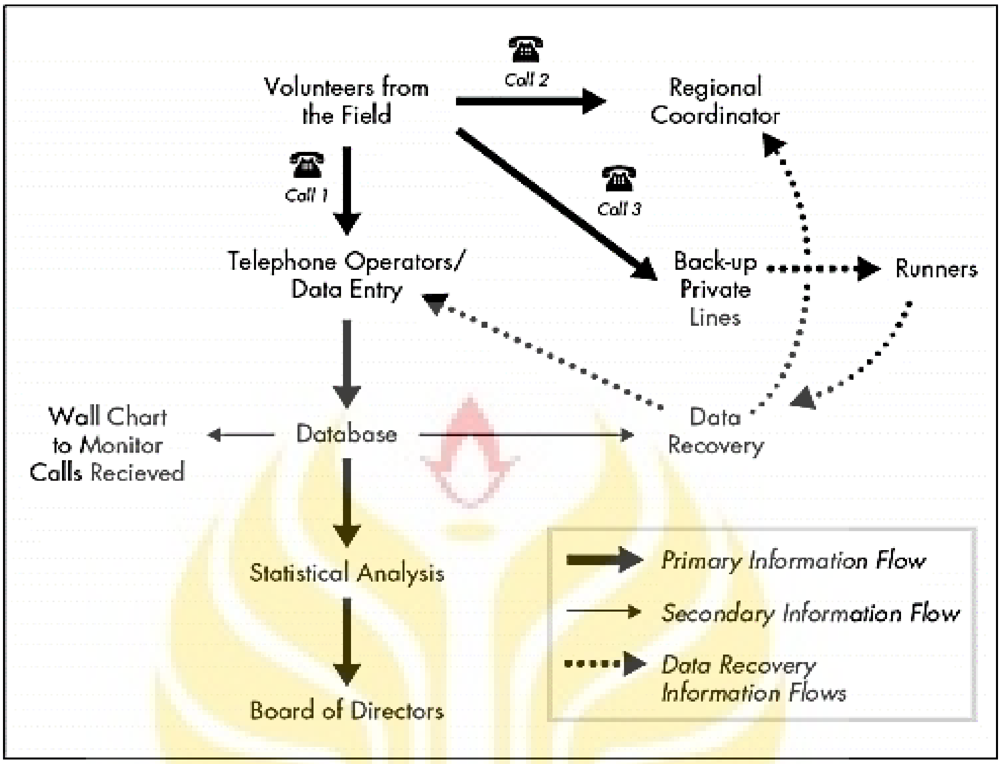
Organisasi *quick count* dapat berbeda antara satu lembaga dengan lembaga lainnya, hal ini lebih disesuaikan dengan kebutuhan masing-masing lembaga penyelenggara *quick count.* Gambar 2.2 berikut menyatakan organisai *quick count* yang dikemukakan oleh Estok *et al*. (2002).



Gambar 2.2 Diagram Organisasi *Quick Count*

2.1.18 Komunikasi Data *Quick Count*

Ukuran lokasi pemantauan (TPS) yang mencapai ribuan dengan melibatkan ribuan orang relawan, tentu bukan pekerjaan sederhana, terutama dalam aspek komunikasi data. Organisasi pelaksana harus menyiapkan perangkat komunikasi data yang terpusat. Arus komunikasi dilakukan dua arah : dari relawan (di lokasi TPS terpantau) untuk pengiriman data lapangan dan dari pusat untuk tujuan pengecekan. Berikut Gambar 2.3 yang menyatakan alur informasi *quick count* yang dikemukakan oleh Estok *et al.* (2002).



Gambar 2.3. Alur Informasi *Quick Count*

Tahapan proses dalam *quick count* secara singkat menurut LSI (2006) adalah

1. Menentukan ukuran sampel TPS yang akan diamati
2. Memilih sampel TPS yang akan diamati secara acak
3. Manajemen data (pengamatan, pencatatan, dan analisis data hasil perhitungan suara)
4. Publikasi hasil *quick count*.

2.2 Penelitian Terdahulu

Pada tahun 2011 Pramita Elfa Diana Santi melakukan penelitian berjudul Penentuan Estimasi Interval dari Distribusi Normal dengan Metode Bayes.

Dengan batasan masalah sebagai berikut : (1) Distribusi sampel yang digunakan adalah distribusi normal univariat. (2) Distribusi prior yang digunakan adalah distribusi normal standar dan distribusi invers gamma. (3) Interval kepercayaan yang disusun antara lain: Interval kepercayaan mean (θ) dengan variansi (σ^2) diketahui untuk distribusi prior normal standar dan Interval kepercayaan variansi (σ^2) dengan mean (θ) diketahui untuk distribusi prior invers gamma. Dalam penelitian Pramita Elfa Diana Santi mampu merumuskan distribusi posterior dari distribusi prior normal standar dan distribusi prior invers gamma, merumuskan interval kepercayaan mean (θ) dengan variansi (σ^2) diketahui untuk distribusi prior normal standar, dan merumuskan interval kepercayaan variansi (σ^2) dengan mean (θ) diketahui untuk distribusi prior invers gamma.

Dari penelitian terdahulu tersebut, penulis dirasa perlu untuk meneliti bagaimana analisis estimasi interval dalam inferensi statistik distribusi Binomial dengan metode Bayes sebagai bahan materi mencari nilai proporsi hasil *quick count* berdasarkan sumber pustaka yang ada.

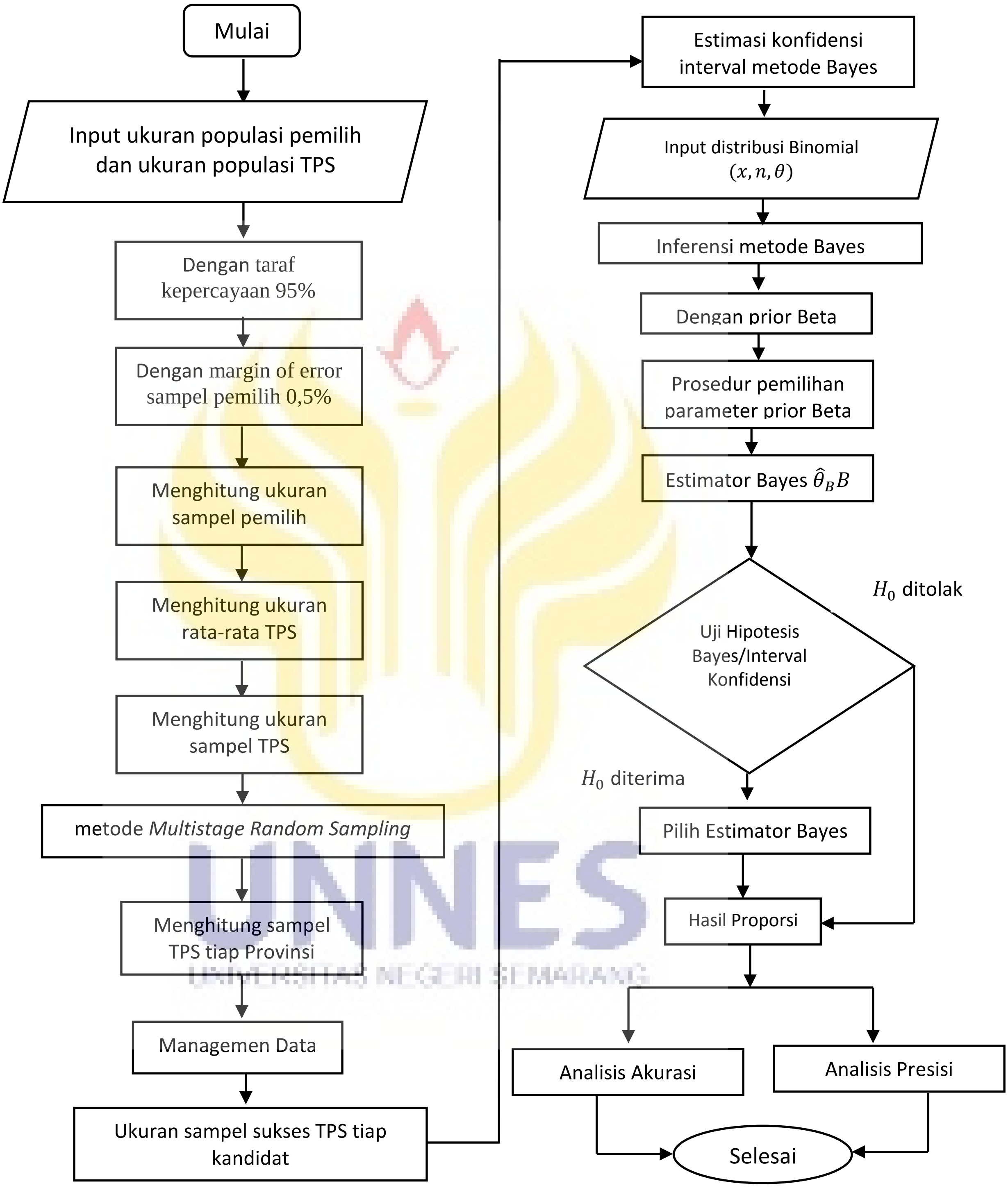
2.3 Kerangka Berpikir

Sebagai kerangka berpikir penulis untuk menjawab permasalahan yang telah dirumuskan dan akan dijabarkan dalam pembahasan tugas akhir ini adalah mengumpulkan informasi tentang ukuran populasi pemilih tetap dan ukuran populasi TPS per Provinsi di Indonesia dari Komisi Pemilihan Umum (KPU) yang kemudian dicari ukuran sampel pemilih dengan taraf kepercayaan dan *margin of error* yang dikehendaki. Berdasarkan ukuran sampel pemilih yang diperoleh tersebut maka ukuran rata-rata TPS di semua Provinsi di Indonesia yang akan digunakan untuk mencari ukuran sampel TPS dapat ditentukan. Lembaga penyelenggara *quick count* menerapkan prinsip probabilitas dalam penarikan sampel. Dalam pengambilan sampel, lembaga penyelenggara *quick count* menggunakan metode *multistage random sampling*. Dengan teknik tersebut dimungkinkan setiap anggota populasi mempunyai peluang yang sama untuk dipilih atau tidak dipilih menjadi responden, sehingga pengukuran pendapat dapat dilakukan dengan tanpa melibatkan semua anggota populasi, hasil survei dapat digeneralisasikan sebagai representasi populasi.

Dalam hal managemen data, LSI melakukan komunikasi data, pencatatan, dan pengolahan data hasil perhitungan cepat (*quick count*). Lembaga penyelenggara menugaskan *surveyor* mengumpulkan data dan dikirim ke pusat untuk dilakukan analisis. Setelah data masuk ke database pusat, maka akan diperoleh ukuran sampel TPS dan ukuran sampel TPS sukses tiap kandidat, perhitungan proporsi dapat dihitung dengan mengestimasi konfidensi interval menggunakan metode Bayes. Metode Bayes memandang parameter sebagai variabel yang menggambarkan pengetahuan awal tentang parameter sebelum pengamatan dilakukan dan dinyatakan dalam suatu distribusi yang disebut dengan distribusi prior dikombinasikan dengan informasi dengan data sampel melalui teorema Bayes, dan hasilnya dinyatakan dalam bentuk distribusi yang disebut distribusi posterior yang selanjutnya menjadi dasar untuk inferensi di dalam metode Bayes. Setelah distribusi posterior terbentuk, maka dapat diperoleh estimasi titik, interval, dan uji hipotesis Bayes untuk parameter yang diestimasi.

Setelah hasil proporsi dengan mengestimasi konfidensi interval menggunakan inferensi statistik metode Bayes diperoleh, selanjutnya adalah menganalisis akurasi dan presisi hasil proporsi *quick count*. Analisis akurasi dan presisi dapat dilakukan dengan cara membandingkan hasil *quick count* oleh LSI dengan hasil perolehan suara resmi KPU.

Kerangka berpikir penulis menganalisis *quick count* metode *Multistage Random Sampling* dengan estimasi konfidensi interval menggunakan metode Bayes dengan studi kasus *quick count* Pemilihan Presiden tahun 2014 oleh LSI dijelaskan dalam bentuk bagan berikut.



Gambar 2.4 Bagan Kerangka Berpikir

BAB V

PENUTUP

5.1 Simpulan

Dari permasalahan pada bab sebelumnya, dapat disimpulkan beberapa hal sebagai berikut.

1. Prosedur analisis perhitungan cepat (*quick count*) metode *Multistage Random Sampling* dengan estimasi konfidensi intervalnya menggunakan perhitungan inferensi statistik yaitu mencari proporsi ukuran sampel mewakili populasi menggunakan metode Bayes adalah sebagai berikut.

   1) Mencari ukuran populasi pemilih dan ukuran populasi TPS.
   2) Menghitung ukuran sampel pemilih dengan perumusan

   $$Ukuran\ Sampel\ Pemilih = \frac{NZ^2[p(1-p)]}{NE^2 + Z^2[p(1-p)]}$$

   3) Menghitung ukuran sampel TPS

   Ukuran sampel TPS yang akan menjadi unit analisis dalam *quick count* ini dapat ditentukan dengan menghitung terlebih dahulu ukuran rata-rata TPS dengan rumus

   $$Ukuran\ rata-rata\ TPS = \frac{Ukuran\ Populasi\ Pemilih}{Ukuran\ Populasi\ TPS}$$

   Setelah ukuran rata-rata TPS diperoleh, ukuran sampel TPS dapat dihitung dengan rumus

   $$Ukuran\ sampel\ TPS = \frac{Ukuran\ Sampel\ Pemilih}{Ukuran\ Rata-rata\ \ TPS}.$$

4) Memilih sampel TPS dengan metode *Multistage Random Sampling*

Banyaknya proses pengambilan sampel TPS dengan *metode Multistage Random Sampling* tergantung pada berapa ukuran sampel TPS yang dibutuhkan. Ukuran sampel TPS yang akan diambil pada masing-masing Provinsi adalah proporsional, berbanding lurus dengan ukuran populasi pemilih dan ukuran populasi TPS yang terdapat di Provinsi tersebut. Semakin banyak ukuran populasi pemilih dan ukuran populasi TPS, maka semakin banyak pula ukuram sampel TPS yang akan diambil dari Provinsi tersebut. Rumus yang digunakan adalah

$$n_i = n \times P_i$$

Kemudian akan dipilih wilayah administrasi yang lebih kecil, yaitu Kabupaten/Kota secara acak.

5) Proses Managemen Data

Metode *Multistage Random Sampling* hanya merupakan metode pemilihan TPS tersampel. Analisis statistikanya menggunakan metode Inferensi Statistik, yaitu berupa perhitungan proporsi perolehan suara pemilih dari seluruh sampel TPS, untuk mengetahui penyebaran suara untuk masing-masing kandidat Pemilihan Umum. Prosedur estimasi konfidensi interval menggunakan metode Bayes sebagai berikut.

a. Memilih parameter prior beta
b. Mencari Distribusi Posterior dan estimator bayes
c. Estimasi Interval Konfidensi Bayes dan Uji Hipotesis

Adapun analisis Statistika Deskripstif, yaitu

$$P_A = \frac{\sum_{i=1}^{n} X_A}{n} \text{ dengan } X_A = \begin{cases} 1, & \text{jika memilih kandidat A} \\ 0, & \text{jika tidak memilih kandidat B} \end{cases}$$

Demikian metode perhitungan proporsi perolehan suara pemilih untuk setiap calon peserta Pemilu lainnya.

2. Analisis akurasi dan presisi dalam perhitungan cepat (quick count) dapat dilakukan dengan cara membandingkan hasil *quick count* yang telah diperoleh dengan hasil perhitungan resmi penyelenggara pemilihan umum yaitu KPU. Hasil *quick count* memiliki akurasi yang tinggi jika *quick count* tersebut dapat meramalkan siapa pemenang dan urutan komposisi pemenang pemilu. Sedangkan *quick count* dikatakan memiliki presisi yang tinggi jika memiliki selisih proporsi yang kecil untuk masing-masing kandidat antara hasil *quick count* dengan hasil perhitungan akhir penyelenggara pemilu yaitu KPU. Suatu *quick count* dapat dikatakan berhasil jika memilik selisih hasil perhitungan yang lebih kecil daripada tingkat ketelitian yang ditoleransi (*margin of error*).

   Perhitungan cepat *(quick count*) yang dilakukan oleh Lembaga Survei Indonesia (LSI) pada Pemilu Presiden tahun 2014 metode *Multistage Random Sampling* dengan estimasi konfidensi interval menggunakan metode Bayes yang telah dibandingkan dengan perolehan hasil perhitungan resmi Komisi Pemilihan Umum (KPU) diketahui bahwa :

   1) urutan perolehan suara untuk masing-masing pasangan calon Presiden dan Wakil Presiden adalah sama antara hasil perhitungan resmi yang dikeluarkan oleh Komisi Pemilihan Umum (KPU) dan perhitungan

cepat (*quick count*) yang dikeluarkan oleh Lembaga Survei Indonesia (LSI). Jadi Lembaga Survei Indonesia (LSI) pada Pemilihan Umum Presiden tahun 2014 memiliki tingkat akurasi yang tinggi.

2) selisih perolehan suara untuk masing-masing pasangan calon Presiden dan Wakil Presiden tahun 2014 antara hasil perhitungan resmi yang dikeluarkan oleh Komisi Pemilihan Umum (KPU) dan perhitungan cepat (*quick count*) yang dilaksanakan oleh Lembaga Survei Indonesia (LSI) adalah 0,17%. Asumsikan perbedaan ini cukup kecil sekali, masih terletak dalam batas kesalahan yang ditoleransi (*margin of error*) atau dengan kata lain masih berada di bawah 0,62%. Jadi secara presisi, *quick count* dengan menggunakan metode *Multistage Random Sampling* oleh Lembaga Survei Indonesia (LSI) juga memiliki presisi yang tinggi.

5.2 Saran

1. Dalam praktiknya, organisasi peneliti melakukan perhitungan cepat (*quick count*) dalam Pemilu Presiden maupun Pemilu Daerah menggunakan *Software* yang telah dirancang sedemikian rupa yang bersifat rahasia dan praktis dalam penggunaannya, namun perlu diketahui dan dipahami pula cara perhitungan secara manualnya.
2. Perlu adanya analisis *quick count* dengan perumusan masalah yang lebih kompleks sesuai kebutuhan masyarakat akan fenomena aksi saling klaim

kemenangan pada Pemilu Presiden tahun 2014 yang dianalisis dengan ilmu statistika yang lebih modern.

DAFTAR PUSTAKA

Bain, L.J. & Engelhardt, M. 1992. *Introduction to Probability and Mathematical Statistics*. Second Edition. California: Duxbury Press.

Berger, C. 1990. *Statistical Inference*. New York: Pasific Grove.

Bolstad, W. M. 2007. *Introduction to Bayesian Statistics Second Edition*. America: A john Wiley & Sons, Inc.

Box, G.E.P. & Tiao, G.C. 1973. *Bayesin Inference In Statistical Analysis*. Philippines: Addision-Wesley Publishing Company, Inc.

Chandra, S.A. 2011. *Inferensi Statistik Distribusi Binomial dengan Metode Bayes Menggunakan Prior Konjugat*. Program Studi Statistika Jurusan Matematika dan Ilmu Pengetahuan Alam Universitas Diponegoro; Semarang.

Cochran, W.G. 1977. *Sampling Techniques Third Edition*. America: A john Wiley & Sons, Inc.

Dajan, Anto. 1984. *Pengantar Metode Statistika Jilid 2*. Jakarta; LP3ES.

Elfa, P.D.S. 2009. Skripsi. *Penentuan Estimasi Interval dari Distribusi Normal dengan metode bayes.* Program Studi Statistika Jurusan Matematika dan Ilmu Pengetahuan Alam Universitas Diponegoro; Semarang.

Eriyanto. 1990. *Metodelogi Polling Memberdayakan Suara Rakyat*. Bandung: PT Remaja Rosdakarya.

Estok M, Nevitte N dan Cowan G. 2002. *The Quick Count and Election Observation*. Washington: NDI.

Gilliland, D dan Melfi, V. 2010. A Note on Confidence Interval Estimation and Margin of Error. *Journal of Statistics Education*. Vol 18, Number 1

Komisi Pemilihan Umum (KPU). 2014. *Pemilihan Presiden 2014*. [Terhubung berkala]. http://www.kpu.go.id/ [28 Desember 2015].

Kurnia, Ahmad. 2015. *Managemen Penelitian: Teknik Sampling*. Jakarta: Reconiascript Publishing.

LSI. 2006. *Panduan Menyelenggarakan Quick Count*. [Terhubung berkala]. http://www.20julbooklsi.pdf [27 Desember 2015].

Lembaga Survei Indonesia (LSI) dan Saiful Munjani Research & Consulting (SMRC). 2014. *Laporan Quick Count Pemilihan Presiden 9 Juli 2014*. [Terhubung berkala]. http://www.kpu.go.id/koleksigambar/LSI-SMRC_Laporan_Quick_Count_Pemilu_Presiden_2014.pdf [20 November 2015].

Margono, S. 2004. *Metodologi Penelitian Pendidikan*. Jakarta: Rineka Cipta.

Scheaffer RL, Mendenhall W dan Ott L. 1990. *Elementary Survey Sampling*. Boston: PWS-Kent.

Soejoeti, Z dan Soebanar. 1988. *Inferensi Bayesian*. Jakarta: Karunika Universitas Terbuka.

Solimun. 2014. Estimating Confidence Interval of Mean Using Classical, Bayesian, and Bootstrap Appoaches. *International Journal of Mathematical Analysis*. Vol 8 Number 48 Hal. 2375-2383.

Sudjana. 2005. *Metoda Statistika*. Bandung: Tarsito.

Supranto, J. 1992. *Teknik Sampling Untuk Survei dan Eksperimen*. Jakarta: PT Rineka Cipta.

Ujiyati T.P. 2004. *Quick Count*. [Terhubung berkala]. http://www.lp3es.or.id/program/pemilu2004/QCount.htm [19 November 2015].

Walpole, R.E dan Myers, R.H. 1995. *Ilmu Peluang dan Statistika untuk Insinyur dan Ilmuwan Edisi ke-4*. Alih bahasa oleh Sembiring, R.K. Bandung: ITB.