

BAB III

METODOLOGI PENELITIAN

A. Metoda, Waktu dan Tempat Penelitian

Pada Bab III ini akan dibahas berbagai aspek yang berkaitan dengan metodologi penelitian seperti metoda penelitian, waktu dan tempat penelitian, sampel penelitian, instrument penelitian, variabel penelitian, definisi operasional variabel penelitian, alur penelitian, teknik analisis data, dan hipotesis penelitian.

1. Metoda Penelitian

Metoda penelitian yang digunakan dalam penelitian ini adalah metoda deskriptif kuantitatif. Data kuantitatif yang digunakan dalam penelitian ini adalah respon peserta UN bahasa Inggris SMA se-Kabupaten Purwakarta tahun pelajaran 2012/2013 dengan *barcode* soal USB-E08, program IPS. Selain itu, data kuantitatif penelitian juga berasal dari panelist berupa hasil *judgment* terhadap butir soal UN dengan menggunakan metoda *standard setting* dari Nedelsky dan Angoff. Hasil *judgment* panelist tersebut berupa *cut off score* yang kemudian digunakan untuk menentukan kelulusan peserta tes dalam UN bahasa Inggris.

2. Waktu dan Tempat Penelitian

Pengumpulan data penelitian dilakukan pada awal bulan Maret 2014. Data penelitian yang pertama berupa respon peserta tes berasal dari Pusat Penilaian Pendidikan (PUSPENDIK) Kementerian Pendidikan Nasional Republik Indonesia, Jakarta. Kegiatan selanjutnya dilakukan pada minggu ke 1 bulan April 2014, yaitu pengambilan data kuantitatif melalui diskusi panel dengan guru-guru SMA yang tergabung dalam Musyawarah Guru Mata Pelajaran (MGMP) bahasa Inggris Kabupaten Purwakarta yang diselenggarakan di SMA Negeri 2 Purwakarta sebagai sekretariat kegiatan MGMP, yang kemudian beberapa guru tersebut bertindak sebagai panelis. Kegiatan diskusi panel tersebut berupa pengarahan dan petunjuk teknis pengisian lembaran *judgment* untuk menentukan *standard setting*

Engkos Kosasih, 2014

**KOMPARASI METODA NEDELSKY DAN ANGOFF DALAM PENETAPAN STANDARD SETTING
KELULUSAN UJIAN NASIONAL MATA PELAJARAN BAHASA INGGRIS SMA**

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

dengan pendekatan metoda Angoff dan Nedelsky. Untuk metoda Angoff, panelis menganalisis tingkat kesulitan (*difficulty index*) butir soal dan memberikan estimasi berupa proporsi persentase soal yang dijawab benar oleh peserta tes. Sedangkan untuk metoda Nedelsky, panelist menganalisis keberfungsian pengecoh (*distractor*) pada opsi butir soal dan untuk kemudian memberikan estimasi terhadap opsi tiap butir soal yang dieliminasi atau opsi yang memiliki probabilitas tidak akan dipilih oleh peserta tes.

Kegiatan berikutnya dilakukan antara kurun waktu bulan April sampai dengan Mei, yaitu pengolahan dan analisis data sampai kepada tahap akhir pengolahan dan analisis data, yaitu menggeneralisasi temuan-temuan dalam penelitian untuk menjawab hipotesis penelitian yang telah diajukan.

B. Instrumen Penelitian

Instrumen yang digunakan untuk penelitian ini adalah sebagai berikut :

1. Data set yaitu berupa respon peserta UN bahasa Inggris SMA tahun pelajaran 2012/2013 Program IPS dengan barcode USB-E08 dengan jumlah butir soal sebanyak 50 butir dalam tes objektif pilihan ganda (*multiple choice*);
2. Format estimasi yang digunakan oleh panelist dalam menentukan *cut off score* kelulusan untuk kedua metoda standard setting Angoff dan Nedelsky;
3. Format rekapitulasi hasil estimasi panelist dalam menentukan *cut off score* untuk kedua *standard setting* yang digunakan.

C. Populasi dan Sampel Penelitian

Populasi yang digunakan sebagai sumber data dalam penelitian merujuk kepada suatu kelompok subjek yang dipilih untuk mewakili seluruh anggota kelompok. Hal tersebut sejalan dengan pendapat Ali (2011: 82) yang mendefinisikan populasi sebagai berikut:

‘Dalam metodologi riset, kelompok besar subjek riset disebut dengan populasi subjek atau populasi riset, sedangkan bagian dari kelompok yang

mewakili kelompok besar itu disebut dengan sampel subjek atau sampel riset?

Sejalan dengan pendapat Ali tersebut, populasi dan sampel yang digunakan dalam penelitian ini dideskripsikan sebagai berikut:

1. Populasi

Populasi yang digunakan sebagai sumber data penelitian adalah guru bahasa Inggris SMA yang tergabung dalam MGMP bahasa Inggris se-Kabupaten Purwakarta yang bertindak sebagai panelis.

2. Sampel

Sampel yang mewakili populasi tersebut adalah guru bahasa Inggris dengan jumlah 25 panelist.

Ada beberapa teknik penyampelan yang digunakan untuk tujuan penelitian. Ali (2011: 102) mengklasifikasikan teknik pensampelan sebagai berikut:

‘Secara umum teknik penyampelan dapat dibedakan kedalam dua kategori utama, yaitu 1) penyampelan berpeluang (*probability sampling*) dan 2) penyampelan tidak berpeluang (*non-probability sampling*). Penyampelan berpeluang dilakukan dengan teknik-teknik yang memungkinkan setiap subyek mempunyai peluang yang sama untuk terpilih sebagai anggota sampel. Teknik-teknik yang termasuk ke dalam kategori ini adalah 1) penyampelan random, 2) penyampelan stratifikasi, 3) penyampelan klaster, 4) penyampelan berjenjang. Adapun yang termasuk kategori penyampelan tak berpeluang adalah 1) penyampelan kuota, 2) penyampelan purposif, dan 3) penyampelan aksidental’

Berdasarkan pendapat Ali tersebut, teknik penyampelan yang digunakan untuk tujuan penelitian ini adalah teknik penyampelan stratifikasi. Penulis mengidentifikasi karakteristik umum populasi, kemudian mengelompokkan unit-unit populasi yang berkarakteristik sama dalam berbagai strata untuk kemudian masing-masing strata tersebut dipilih secara random. Karakteristik umum populasi tersebut adalah sebagai berikut:

1. Guru bahasa Inggris dengan kualifikasi pendidikan sarjana pendidikan bahasa Inggris;
2. Telah mengajar dibidangnya sedikitnya lima tahun;

3. Telah memiliki sertifikasi kependidikan.

Kemudian yang menjadi responden adalah peserta UN SMA program IPS tahun ajaran 2012/2013 dengan barcode soal USB-E08 se-Kabupaten Purwakarta sebanyak 134 siswa.

D. Variabel Penelitian

Ada dua variabel bebas dalam penelitian ini, yaitu *cut off score* dari *standard setting* dengan metoda Angoff dan metoda Nedelsky. Sedangkan variabel terikatnya adalah skor baku peserta tes UN. Menurut pendapat Ali (2011:71) menjelaskan bahwa, “Variabel bebas merupakan variabel yang kemunculannya diasumsikan menjadi sebab munculnya variabel terikat. Variabel terikat merupakan variabel yang kemunculannya diasumsikan sebagai akibat dari adanya variabel bebas.

E. Definisi Operasional

Definisi operasional variabel yang digunakan dalam penelitian ini adalah sebagai berikut:

a. Skor Baku

Susetyo (2011; 37) menjelaskan bahwa yang dimaksud dengan skor baku adalah sebagai berikut:

‘Dari data dapat dibentuk data baru yang diperoleh dari penyimpangan data dari rata-rata yang dinyatakan dalam satuan simpangan baku dan bilangan tersebut dinamakan dengan bilangan baku atau skor baku dan dilambangkan dengan T’

Pendapat lainnya tentang skor baku juga dikemukakan oleh Naga (2013;117). Dali menjelaskan bahwa, “Nilai baku adalah letak nilai (jarak) terhadap rerata yang dinyatakan dengan satuan simpangan baku yakni simpangan baku jaraknya dari rerata”.

Dari kedua pendapat tersebut di atas dapat disimpulkan bahwa skor baku adalah skor yang ditunjukkan dengan suatu skala dan menggambarkan perbandingan satu individu dengan individu lain dalam kelompok.

b. *Cut off score*

Istilah *cut off score* (*cutscore*, *cut off score*, *mastery score*, *passing grade*, *passing score*, *passing criterion*) merupakan istilah umum yang digunakan dalam standard setting ketuntasan/kelulusan. Zieky & Perie (2005: 2) menjelaskan bahwa;

'Cut scores are selected points on the score scale of a test. The points are used to determine whether a particular test score is sufficient for some purpose. For example, student performance on a test may be classified into one of several categories such as basic, proficient, or advanced on the basis of cut scores'.

Dari penjelasan yang dikemukakan oleh Zieky & Perie tersebut dapat disimpulkan bahwa *cut off score* adalah poin tertentu yang digunakan pada skala skor hasil test yang digunakan untuk menentukan tingkat penguasaan terhadap kategori yang telah ditetapkan.

Cut off score dari metoda Nedelsky diperoleh dengan cara merata-ratakan skor (*trimmed mean*) dari keseluruhan *judgment* dengan membuang skor yang ekstrim, yaitu skor yang tertinggi dan skor yang terendah sedangkan *cut off score* dari metoda Angoff diperoleh dengan cara merata-ratakan jumlah proporsi tiap butir soal yang dijawab betul oleh peserta tes dari keseluruhan *judgment* yang ditetapkan panelis.

F. Alur Penelitian

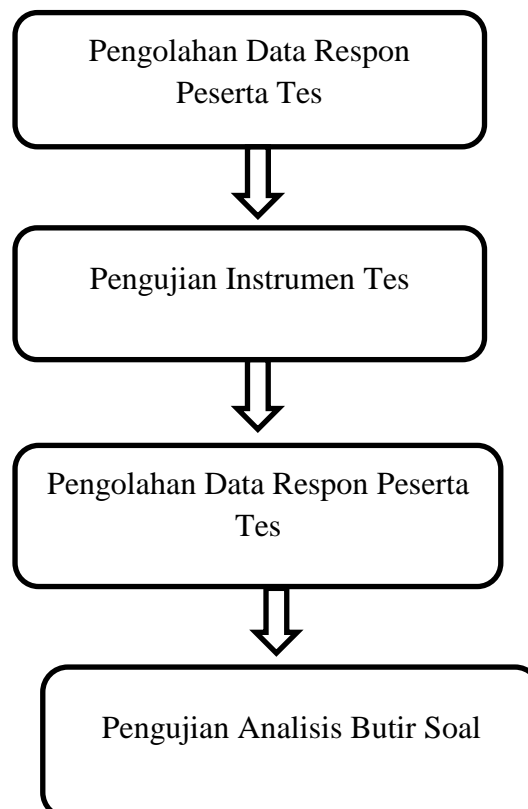
Kegiatan penelitian yang dilakukan melalui tahapan-tahapan berikut ini:

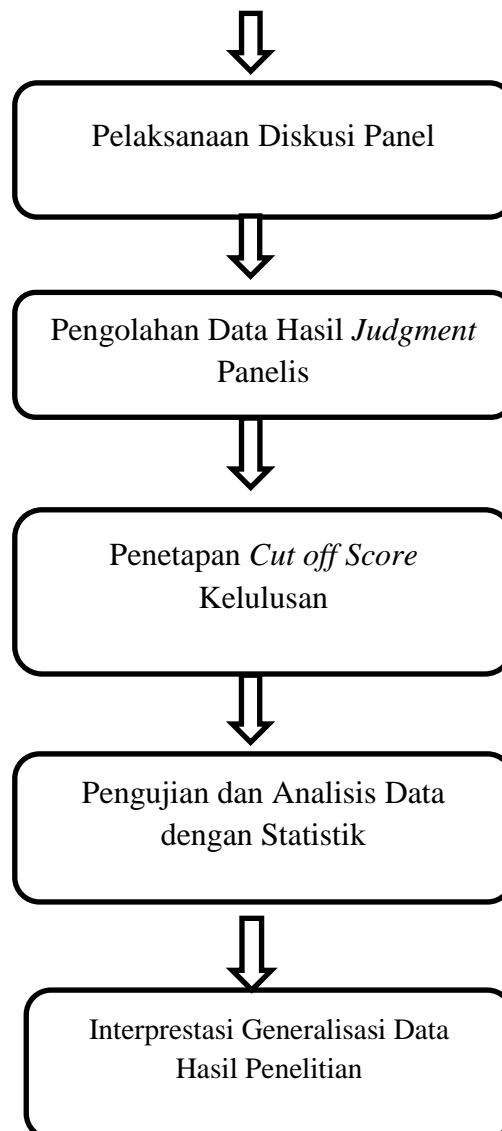
1. Pengolahan data respon peserta tes UN untuk mendapatkan skor mentah (*raw score*) dan untuk kemudian skor mentah tersebut ditransformasikan dengan menggunakan *Z score* untuk mendapatkan skor baku/skor standar.
2. Pengujian terhadap kelayakan instrumen, yaitu pengujian terhadap validitas dan reliabilitas perangkat tes yang digunakan.
3. Pengujian terhadap analisis butir soal yaitu pengujian terhadap a) tingkat kesukaran, b) daya beda, dan c) distraktor.
4. Melaksanakan diskusi panel yang diselenggarakan melalui kegiatan Musyawarah Guru Mata Pelajaran (MGMP) bahasa Inggris. Ada 25

panelist yang memiliki persyaratan kualifikasi akademik dan persyaratan administrasi yang bersedia menjadi panelis. Kepada panelis tersebut dilakukan diskusi mengenai petunjuk pelaksanaan penentuan estimasi *judgment* berdasarkan kedua metoda *standard setting* yang diberikan.

5. Pengolahan data kuantitatif hasil *judgment* panelis.
6. Penentuan *cut off score* berdasarkan hasil *judgment* panelis untuk kedua metoda *standard setting*.
7. Menentukan kelulusan peserta UN dengan cara membandingkan skor baku yang diperoleh peserta tes dengan *cut off score* dari masing-masing metoda *standard setting*.
8. Melakukan pengujian normalitas dan homogenitas data skor panelis.
9. Melakukan pengujian lebih lanjut terhadap data skor panelis.
10. Tahap akhir dari alur penelitian ini adalah menginterpretasi dan menggeneralisasi keseluruhan hasil temuan-temuan penelitian.

Secara garis besar alur kegiatan penelitian ini digambarkan dalam bagan alur seperti di bawah ini:





Gambar 3.1 Bagan Alur Kegiatan Penelitian

G. Teknis Analisis Data

Teknis analisis data dalam penelitian ini dilakukan melalui beberapa tahapan, yaitu:

1. Pengujian persyaratan kelayakan perangkat tes. Uji kelayakan yang dilakukan adalah pengujian terhadap validitas dan reliabilitas perangkat tes. Untuk pengujian validitas dilakukan terhadap validitas isi (*content*

validity) yaitu melalui uji kecocokan butir soal dengan indikator pada Standar Kompetensi Kelulusan (SKL).

2. Kemudian uji reliabilitas perangkat tes dengan menggunakan perhitungan reliabilitas *Cronbach Alpha*. Penggunaan *Cronbach Alpha* dikarenakan skor yang akan didapat berupa skor dikotomi. Berikut persamaan yang digunakan untuk menghitung koefisien reliabilitas (Juliandi, 2014: 2)

$$r = \left[\frac{k}{(k-1)} \right] \left[1 - \frac{\sum \sigma_b^2}{\sigma_t^2} \right]$$

Keterangan:

r = koefisien reliabilitas instrument (*cronbach alpha*)

k = banyaknya butir pertanyaan atau banyaknya soal

$\sum \sigma_b^2$ = total varians butir

σ_t^2 = total varians

Untuk dijadikan pegangan, para pakar terkait biasanya hanya menyarankan agar pelaku riset menghindari penggunaan instrumen yang memiliki derajat kereliabelan rendah. Secara kuantitatif, derajat kereliabelan dianggap rendah apabila indeks yang diperoleh dari hasil pengujian adalah sama dengan atau lebih dari 0.40 ($r \leq 0,40$) Ali (2011: 141).

3. Pengolahan data hasil respon peserta tes. Kegiatan menginput data hasil respon peserta tes dimaksudkan mengolah data tersebut untuk mendapatkan skor mentah (*raw score*). Data mentah tersebut kemudian diolah dengan cara mencari standar deviasinya dan ditransformasikan kedalam *z score* untuk mendapatkan skor baku/standar. Untuk mendapatkan skor mentah (*raw score*), peneliti menggunakan rumus sebagai berikut.

$$\text{Nilai (X)} = \frac{\text{Skor Mentah}}{\text{Skor Maksimum}} \times 100$$

Dan kemudian untuk mendapatkan skor baku, peneliti menggunakan rumus sebagai berikut;

$$Z = \frac{X - \bar{x}}{S} \quad (\text{Susetyo, 2011: 37})$$

dimana:

Z = z skor, X = skor \bar{x} = rata-rata, S = standar deviasi atau simpangan baku

Untuk menghindari skor yang bertanda negatif maka digunakan T skor dimana T skor menggunakan rata-rata X = 50, dan simpangan baku, S = 10, dengan rumus;

$$\text{Skor baku} = 50 + 10z \quad (\text{Susetyo, 2011: 38})$$

4. Melakukan analisis perangkat tes seperti uji tingkat kesukaran soal (*difficulty level*), daya beda (*item discrimination*), dan keberfungsian distractor (*item functioning respon*) dengan menggunakan teori tes klasik (*classical test theory*).
 - a. Untuk menghitung tingkat kesukaran soal (*difficulty level*), peneliti menggunakan rumus sebagai berikut:

$$P_i = 0.5 (P_{Ti} + P_{Ri}), \quad (\text{Susetyo, 2011: 156})$$

Dimana;

$$P_{Ti} = \frac{f_i(x=1)}{M}$$

f_i = peserta yang menjawab benar

P_{Ti} = taraf kesukaran butir tes untuk kelompok tinggi ke-I

M = jumlah peserta test

$$P_{Ri} = \frac{f_i(x=1)}{M}$$

f_i = peserta yang menjawab benar

P_{Ri} = taraf kesukaran butir tes untuk kelompok rendah ke-i

M = jumlah peserta tes

Pembagian tingkat kesukaran menurut Witherington seperti yang dikutip Susetyo (2011: 154) adalah sebagai berikut:

Tabel 3.1 Tingkat Kesukaran Butir Soal

Rentang	Tingkat Kesukaran
$0,00 \leq P \leq 0,24$	Sukar
$0,25 \leq P \leq 0,74$	Sedang
$0,75 \leq P \leq 1,00$	Mudah

- b. Untuk menghitung daya beda butir soal (*item discrimination*) digunakan rumus;

$$DP = \frac{2(KA-KB)}{n} \quad (\text{Puspendik, 2008: 4})$$

dimana,

DP : daya pembeda

KA : banyak siswa pada kelompok atas yang menjawab benar

KB : banyak siswa pada kelompok bawah yang menjawab benar

N : banyaknya siswa

Pembagian daya beda menurut Ebel R. (1979: 267) seperti yang dikutip Susetyo (2011: 161) sebagai berikut:

Tabel 3.2 Indeks Daya Beda

Indeks Daya Beda	Keterangan
$0,70 \leq D \leq 1,0$	Butir memiliki daya beda baik sekali
$0,40 \leq D \leq 0,69$	Butir memiliki daya beda cukup baik
$0,30 \leq D \leq 0,39$	Butir memerlukan revisi sedikit atau tidak
$0,20 \leq D \leq 0,29$	Butir memerlukan revisi atau disisihkan
$0,00 \leq D \leq 0,19$	Butir direvisi total atau disisihkan

- c. Untuk menghitung keberfungsian distraktor peneliti menggunakan rumus;

$$p_{x_i} = \frac{f_{x_i}}{M} \times 100\% \quad (\text{Susetyo, 2011: 170})$$

dimana,

M = jumlah responden

p_{x_i} = proporsi masing-masing pilihan jawaban suatu butir tes

f_{x_i} = frekwensi masing-masing pilihan jawaban suatu butir tes.

Susetyo (2011: 172) menjelaskan bahwa suatu distraktor berjalan dengan baik apabila dapat menggiring peserta tes untuk memilihnya dan sekurang-kurangnya telah dipilih oleh peserta minimal 5% dari jumlah keseluruhan peserta tes.

5. Pengolahan data hasil *judgment* panelis akan berupa *cut off score* kelulusan. Penetapan *cut off score* dengan metoda Nedelsky digunakan rumus;

$$\text{trimmed mean} = \frac{\bar{x}}{N}$$

dimana:

\bar{x} = jumlah skor

N = jumlah panelist

Untuk penghitungan dengan cara *trimmed mean*, skor yang ekstrim yaitu skor tertinggi dan skor terendah tidak dihitung atau dieliminasi.

Sedangkan untuk *cut off score* metoda Angoff digunakan rumus;

$$\frac{x_1 + x_2 + x_3 + x_4 + \dots, x_n}{N}$$

dimana:

x = rata rata skor estimasi panelis

N = jumlah panelist

6. Melakukan uji validasi internal terhadap hasil *judgment* panelis, melalui:
- analisis konsistensi metoda dengan cara mengestimasi *standar error of judgment* (SE_j) dengan rumus:

$$SD / \sqrt{n}$$

dimana:

SD = standar deviasi

n = jumlah panelis

dengan kriteria harga SE_j harus sama atau lebih kecil dari pada SEM

- Analisis terhadap konsistensi *inter-judge* untuk mengetahui kelayakan metoda yang digunakan. Analisis dilakukan dengan cara melihat simpangan baku *cut off score* dari masing-masing metoda. Kriteria pengujian terhadap konsistensi *inter-judge* adalah bahwa kelayakan suatu metoda dapat dilihat dari nilai simpangan bakunya. Apabila simpangan baku yang diperoleh semakin kecil maka semakin layak metoda tersebut digunakan.
- Analisis terhadap signifikansi perbedaan *cut off score* antara dua metoda dengan menggunakan rumus sebagai berikut:

$$| CX - CY | / SEM > 2.77.$$

dimana:

CX = *cut off score* dari metoda Nedelsky

CY = *cut off score* dari metoda Angoff

7. Mentabulasikan data hasil pengolahan respon peserta tes yaitu berupa skor baku dengan *cut off score* dari masing-masing metoda *standard setting*;
8. Pengujian data analisis lebih lanjut yaitu dengan menggunakan uji statitika sebagai berikut:
- Uji normalitas data dilakukan untuk mengetahui apakah data berdistribusi normal atau tidak. Pengujian normalitas data dilakukan dengan menggunakan formula dari statistik uji Lilliefors (*Kolmogorov-Smirnov*).

- b. Uji statistika berikutnya adalah uji homogenitas varian (*uji Levine*). Untuk menguji apakah ragam populasi dari dua sampel diasumsikan homogen atau tidak maka dilakukan uji homogenitas. Untuk itu, asumsi homogenitas ragam populasi dari 2 sampel ini perlu diuji terlebih dahulu. Semua pengujian dilakukan dengan taraf kepercayaan 5% ($\alpha = 0.05$)
- c. Apabila diasumsikan bahwa ragam populasi dari kedua sampel adalah homogen maka langkah selanjutnya dilakukan pengujian statistika dengan *uji-t* 2 sampel independen. Uji-t (*t-test*) yang digunakan adalah untuk sampel berpasangan dimana sampel berasal dari subjek yang sama. Untuk pengujian hipotesis dengan *uji-t* tersebut ditetapkan hipotesis sebagai berikut:

Hipotesis statistik:

$$H_0 : \mu A = \mu B$$

$$H_A : \mu A \neq \mu B$$

Hipotesis penelitian:

H_0 = Tidak ada perbedaan antara metoda Nedelsky dan Angoff dalam penetapan jumlah proporsi kelulusan.

H_A = Ada perbedaan antara metoda Nedelsky dan Angoff dalam penetapan jumlah proporsi kelulusan.

Untuk pengujian perbedaan dua rata-rata populasi (σ), $\sigma_1 = \sigma_2 = \sigma$ masing-masing diketahui maka menggunakan rumus sebagai berikut:

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

dimana :

\bar{x}_1 = rata-rata sampel kelompok 1

\bar{x}_2 = rata-rata sampel kelompok 2

σ = simpangan baku populasi perbedaan rata-rata,

Kriteria Pengujian

Taraf nyata atau signifikansi α , untuk pengujian dua pihak adalah: H_0 diterima jika $-z_{1/2(1-\alpha)} < z < z_{1/2(1-\alpha)}$ diperoleh dari daftar distribusi normal baku dengan peluang $1/2(1 - \alpha)$, sebaliknya H_0 ditolak pada harga lainnya. (Susetyo, 2010: 203)