

BAB III

METODOLOGI PENELITIAN

A. Variabel Penelitian

Secara garis besar variabel penelitian mengenai keterbandingan reliabilitas berdasar metode penskoran *number-right score* dengan metode penskoran *correction for guessing* dibagi menjadi dua yaitu:

1. Variabel bebas pada penelitian ini adalah metode penskoran dengan menggunakan tiga metode penskoran yaitu metode penskoran *number-right score* dan metode penskoran *correction for guessing* yaitu metode penskoran *punishment score* dan metode penskoran *reward score*.
2. Variabel terikat pada penelitian ini adalah reliabilitas tes.

B. Metode Penelitian

Penelitian mengenai keterbandingan reliabilitas tes hasil belajar matematika berdasar metode penskoran *number-right score* dan metode penskoran *correction for guessing* ini merupakan penelitian kuantitatif. Metode yang digunakan dalam penelitian ini adalah metode penelitian kuasi-eksperimental (eksperimen semu). Menurut Ali (2011: 284) “Perbedaan utama studi eksperimental dan studi kuasi-eksperimental terletak pada pemilihan subyek sampel secara random dan penugasan subyek secara random ... pada hakekatnya studi kuasi-eksperimental adalah studi eksperimental”. Jadi metode Penelitian kuasi-eksperimental adalah metode penelitian yang dilakukan melalui kegiatan ekperimentasi atau percobaan yang sengaja dilakukan dan terkontrol. Peneliti tidak melakukan penugasan random saat melakukan penelitian yaitu saat melakukan penskoran menggunakan tiga metode penskoran yaitu metode penskoran *number-right score* dan metode penskoran *correction for guessing* yaitu *punishment score* dan *reward score*.

Maria Agustina Amelia, 2014

KETERBANDINGAN RELIABILITAS TES HASIL BELAJAR MATEMATIKA BERDASAR METODE PENSKORAN NUMBER-RIGHT SCORE DAN METODE PENSKORAN CORRECTION FOR GUESSING

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

Desain penelitian

Desain penelitian yang digunakan adalah desain perbandingan kelompok statis (*the static group comparison design*) yaitu penelitian membandingkan tiga kelompok penelitian. Satu kelompok kontrol reliabilitas skor hasil tesnya menggunakan metode penskoran *number-right score*. Sedangkan dua kelompok eksperimen lain reliabilitas skor hasil tesnya menggunakan metode penskoran *correction for guessing* yaitu *punishment score* dan *reward score*.

Tabel 3.1 Desain perbandingan kelompok berdasar metode penskoran

Model Penskoran		
<i>Number-right score</i>	Correction for guessing	
	<i>Punishment score</i>	<i>reward score</i>
$X_{NRS} = \sum_{i=1}^n x_i$ $\Gamma_{NRS1}, \Gamma_{NRS2}, \dots, \Gamma_{NRS30}$ μ_{NRS}	$X_{RS} = \sum_{i=1}^n x_i - \frac{0}{k}$ $\Gamma_{PS1}, \Gamma_{PS2}, \dots, \Gamma_{PS30}$ μ_{PS}	$X_{RS} = \sum_{i=1}^n x_i + \frac{0}{k}$ $\Gamma_{RS1}, \Gamma_{RS2}, \dots, \Gamma_{RS30}$ μ_{RS}

C. Populasi dan Sampel

Populasi penelitian dibedakan menjadi dua bagian yaitu populasi subyek penelitian dan populasi skor. Populasi subjek adalah seluruh siswa Sekolah Dasar di kota Bandung. Populasi skor adalah skor siswa dan skor butir yang datanya diambil dari hasil tes sampel siswa Sekolah Dasar di Bandung.

Sampel subjek dipilih dari seluruh siswa Sekolah Dasar Negeri di kota Bandung yang berjumlah 575 sekolah. Pengambilan sampel menggunakan teknik *random sampling*. Menurut Taniredja dan Mustafidah (2012: 35) “Teknik *random sampling* disebut juga acak, serampangan, tidak pandang bulu/ tidak pilih kasih, objektif, sehingga seluruh elemen populasi mempunyai kesempatan untuk jadi sampel penelitian”. Jadi teknik *random sampling* dilakukan agar semua populasi

Maria Agustina Amelia, 2014

KETERBANDINGAN RELIABILITAS TES HASIL BELAJAR MATEMATIKA BERDASAR METODE PENSKORAN NUMBER-RIGHT SCORE DAN METODE PENSKORAN CORRECTION FOR GUESSING

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

subjek memiliki peluang yang sama untuk dipilih menjadi sampel subjek. Alasan dipilih teknik pengambilan sampel *random sampling* karena pada tingkat sekolah dasar, penerimaan peserta didik tidak didasarkan pada nilai tertentu atau berdasarkan tes masuk jadi diasumsikan bahwa peserta didik ditingkat sekolah dasar memiliki kemampuan yang merata. Selain hal tersebut sampel subjek dipilih karena sudah selesai mengajarkan materi pecahan yang merupakan materi yang akan diujikan dalam instrumen penelitian. Dari 575 Sekolah Dasar Negeri di kota Bandung dipilih secara random lima Sekolah Dasar Negeri sebagai tempat pelaksanaan penelitian.

Sampel siswa kemudian diberikan tes hasil belajar yang sama yaitu mengenai materi pecahan. Dari tes belajar tersebut, jawaban masing-masing peserta didik akan dinilai menggunakan tiga kelompok berdasar tiga metode penskoran yaitu metode *number-right score*, *punishment score* dan *reward score*. Berdasarkan data kelompok populasi skor tersebut, masing-masing kelompok akan diambil data dengan metode pengambilan Sampel Acak Dengan Pengembalian (SADP) sebanyak sepuluh kali jumlah faktor atau variabel yang memberi pengaruh pada hasil tes. Setelah didapatkan data skor yang diinginkan, hitung reliabilitas dari data skor tersebut. Ulangi pengambilan data secara acak dengan pengembalian sehingga didapat data tiga puluh koefisien reliabilitas. Ketiga puluh data koefisien reliabilitas dari masing-masing kelompok skor yang dijadikan data sampel penelitian.

Sampel skor berupa data koefisien reliabilitas dari kelompok skor siswa yang dinilai menggunakan metode penskoran *number-right score*, metode penskoran *punishment score* dan metode penskoran *reward score*.

D. Instrumen dan Teknik Pengumpulan Data

1. Instrumen

Instrumen yang digunakan dalam penelitian adalah tes berbentuk pilihan ganda dengan empat pilihan jawaban. Ruang lingkup materi yang akan diujikan dalam tes tersebut berkaitan dengan materi pecahan kelas V Sekolah Dasar.

Maria Agustina Amelia, 2014

KETERBANDINGAN RELIABILITAS TES HASIL BELAJAR MATEMATIKA BERDASAR METODE PENSKORAN NUMBER-RIGHT SCORE DAN METODE PENSKORAN CORRECTION FOR GUESSING

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

Sebelum digunakan, instrumen diujicobakan dahulu kepada siswa untuk mengetahui validitas dan reliabilitas instrumen. Sesuai dengan pendapat Arikunto (dalam Taniredja, 2012: 41) “instrumen yang baik harus memenuhi dua persyaratan penting yaitu valid dan reliabel”. Soal uji coba tes dibagi menjadi dua tipe dengan masing-masing tipe berisi 20 butir soal. Uji coba diberikan di tiga kelas (± 100 peserta didik) dengan masing-masing kelas diminta mengerjakan dua jenis soal. Selain itu agar didapat instrumen yang baik, masing-masing butir akan diuji tingkat kesukaran, daya beda dan analisis distraktor.

1. Uji Validitas

Pengujian validitas dilakukan dengan menguji validitas isi dari instrumen. Pengujian validitas isi dilakukan dengan meminta lima orang ahli atau praktisi pendidikan di bidang matematika. Format yang digunakan untuk validasi adalah format skala penilaian (*judgement* ahli) bentuk dikotomi dengan skala satu-nol. Nilai 1 = sesuai dan nilai 0 = tidak sesuai

Tabel 3.2 Format skala penilaian judgement ahli

Butir	Penilai					Jumlah sesuai	persentase
	1	2	.	.	n		
1	x_{11}	x_{12}	.	.	x_{1n}	$\sum x_{1n}$	$\frac{\sum x_{1n}}{100} \%$
2	x_{21}	x_{22}	.	.	x_{2n}	$\sum x_{2n}$	$\frac{\sum x_{2n}}{100} \%$
.	x_{31}	x_{32}	.	.	x_{3n}	.	.
.	x_{41}	x_{42}	.	.	x_{4n}	.	.
m	x_{m1}	x_{m2}	.	.	x_{mn}	$\sum x_{mn}$	$\frac{\sum x_{mn}}{100} \%$

Perhitungan validitas isi menggunakan persentase butir yang cocok dengan indikator

Maria Agustina Amelia, 2014

KETERBANDINGAN RELIABILITAS TES HASIL BELAJAR MATEMATIKA BERDASAR METODE PENSKORAN NUMBER-RIGHT SCORE DAN METODE PENSKORAN CORRECTION FOR GUESSING

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

$$\text{persentase} = \frac{\sum x_{mn}}{m} \times 100\% \quad (3.1)$$

dengan

m = jumlah penilai = 5

$\sum x_{mn}$ = jumlah sesuai butir ke- n

Menurut Susetyo (2011: 92) “Butir soal dikatakan valid jika diperoleh persentase lebih besar dari 50% (persentase $\geq 50\%$)”.

2. Uji Keselarasan (Konkordansi)

Selain uji validitas, dilakukan uji keselarasan (konkordansi) antar penilai menggunakan uji Cochran. Uji Cochran dilakukan untuk melihat apakah ada kesepakatan antar penilai saat melakukan validitas isi terhadap instrumen tes.

Hipotesis uji

H_0 : Terdapat kesamaan penilaian yang diberikan saat melakukan validitas isi instrumen tes

H_1 : Tidak terdapat kesamaan penilaian yang diberikan saat melakukan validitas isi instrumen tes

Signifikansi uji

$$Q = \frac{m(m-1) \sum_{i=1}^m (G_i - \bar{G})^2}{m \sum_{i=1}^n L_i - \sum_{i=1}^n L_i^2} \quad (3.2)$$

dengan

m = jumlah penilai

G_i = jumlah nilai “cocok” pada kolom ke- i

\bar{G} = rata-rata nilai G_i

L_i = jumlah nilai “cocok” pada baris ke- i

Kriteria uji

H_0 ditolak jika $Q < \chi^2_{\text{tabel}}$ atau jika $p < \alpha = 0,05$

H_0 diterima jika $Q > \chi^2_{\text{tabel}}$ atau jika $p > \alpha = 0,05$

Maria Agustina Amelia, 2014

KETERBANDINGAN RELIABILITAS TES HASIL BELAJAR MATEMATIKA BERDASAR METODE PENSKORAN NUMBER-RIGHT SCORE DAN METODE PENSKORAN CORRECTION FOR GUESSING

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

Uji Cochran dilakukan dengan menggunakan *software Statistical Package for Social Sciences* (SPSS).

3. Uji Reliabilitas

Setelah diperoleh hasil validitas isi instrumen diujicobakan di lapangan. Hasil uji coba digunakan untuk mengukur reliabilitas tes. Koefisien reliabilitas dihitung menggunakan teknik perhitungan Alpha Cronbach yaitu :

$$\alpha = \frac{k}{k-1} \left(1 - \frac{\sum s_i^2}{s_x^2} \right) \quad (3.4)$$

dengan

α = koefisien korelasi Alpha Cronbach

k = banyaknya butir soal

s_x^2 = varians skor pada keseluruhan tes X

s_i^2 = varians skor pada belahan tes ke- i

Pendapat Naga (2012: 241) mengenai kriteria koefisien reliabilitas Alpha Cronbach :

Tabel 3.3 Kriteria umum koefisien reliabilitas Alpha Cronbach

Reliabilitas Alpha Cronbach	Kriteria
$\rho_\alpha < 0,50$	Tidak dapat diterima
$0,50 \leq \rho_\alpha < 0,60$	Kurang
$0,60 \leq \rho_\alpha < 0,70$	Bermasalah
$0,70 \leq \rho_\alpha < 0,80$	Dapat diterima
$0,80 \leq \rho_\alpha < 0,90$	Baik
$0,90 \leq \rho_\alpha$	Istimewa

Maria Agustina Amelia, 2014

KETERBANDINGAN RELIABILITAS TES HASIL BELAJAR MATEMATIKA BERDASAR METODE PENSKORAN NUMBER-RIGHT SCORE DAN METODE PENSKORAN CORRECTION FOR GUESSING

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

Butir soal dikatakan reliabel jika memiliki koefisien reliabilitas lebih besar dari 0,7 ($\alpha \geq 0,7$).

4. Analisis Tingkat Kesukaran

Tingkat kesukaran merupakan perbandingan jumlah siswa yang menjawab benar suatu butir soal dengan jumlah seluruh peserta tes. Tingkat kesukaran biasa dinyatakan dalam bentuk proporsi dan berkisar dari nilai 0,00 hingga 1,00. Jika nilai tingkat kesukaran makin mendekati nol, maka butir soal semakin sukar. Sebaliknya jika nilai tingkat kesukaran makin mendekati satu, maka butir soal semakin mudah. Berikut ini klasifikasi tingkat kesukaran menurut Witherington (dalam Susetyo, 2011: 154) :

Tabel 3.4 Klasifikasi tingkat kesukaran

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

Butir soal pada konstruksi tes Sumatif seperti Ujian Akhir bab dengan materi pecahan diharapkan memiliki tingkat kesukaran tiap butir dapat bervariasi dari butir dengan tingkat kesukaran mudah hingga sukar. Penyebaran tingkat kesukaran yang berdistribusi normal dengan sebaran soal sukar 20%, soal sedang 50% dan soal mudah 30% dari keseluruhan jumlah soal.

5. Analisis Daya Beda

Daya beda digunakan untuk membedakan siswa yang berkemampuan tinggi dan berkemampuan rendah. Analisis daya beda menggunakan korelasi point biserial yang perhitungannya dilakukan menggunakan program ITEMAN. Pembagian daya beda menurut Ebel (Susetyo, 2011: 161)

Tabel 3.5 Pembagian daya beda

Maria Agustina Amelia, 2014

KETERBANDINGAN RELIABILITAS TES HASIL BELAJAR MATEMATIKA BERDASAR METODE PENSKORAN NUMBER-RIGHT SCORE DAN METODE PENSKORAN CORRECTION FOR GUESSING

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

Indeks daya beda	Klasifikasi
$0,70 \leq D \leq 1,00$	Daya beda baik
$0,40 \leq D \leq 0,69$	Daya beda cukup baik
$0,30 \leq D \leq 0,39$	Perlu sedikit revisi
$0,20 \leq D \leq 0,29$	Perlu revisi atau disisihkan
$0,00 \leq D \leq 0,19$	Revisi total atau disisihkan

Jika indeks daya beda semakin mendekati nilai satu maka butir soal semakin baik dalam membedakan siswa dengan kemampuan tinggi dengan siswa dengan kemampuan rendah. Menurut Nunnally (Dali, 1992: 7) “butir soal dengan daya beda di atas 0,2 sudah cukup baik. Butir soal dengan indeks daya beda kurang dari 0,2 perlu dibuang karena kurang baik dalam membedakan kemampuan siswa”.

6. Analisis Distraktor

Analisis distraktor dilakukan untuk mengetahui apakah distraktor berjalan dengan baik sebagai pengecoh. Distraktor yang baik juga mencegah siswa untuk melakukan tebakan saat menjawab soal tes pilihan ganda (Susetyo, 2011 : 172). Analisis distraktor dilakukan dengan menggunakan proporsi sebagai berikut:

$$P = \frac{f}{N} \times 100\% \quad (3.5)$$

dengan

P = proporsi distraktor

f = frekuensi jawaban siswa

N = Jumlah seluruh siswa

Distraktor dikatakan berfungsi dengan baik jika proporsi masing-masing distraktor adalah 5%.

2. Teknik Pengumpulan data :

Maria Agustina Amelia, 2014

KETERBANDINGAN RELIABILITAS TES HASIL BELAJAR MATEMATIKA BERDASAR METODE PENSKORAN NUMBER-RIGHT SCORE DAN METODE PENSKORAN CORRECTION FOR GUESSING

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

Teknik pengumpulan data dilakukan dengan menggunakan tes hasil belajar matematika dengan materi pecahan untuk siswa Sekolah Dasar kelas V. Butir soal didapatkan dari butir-butir soal yang telah diujicoba dan direvisi. Langkah-langkah pengumpulan data yang dilakukan adalah :

1. Instrumen tes yang sama diberikan kepada sampel siswa. Kemudian hitung nilai akhir tes menggunakan tiga metode penskoran yaitu metode penskoran *number-right score*, metode penskoran *punishment score* dan metode penskoran *reward score*.
2. Dari sampel siswa didapatkan tiga kelompok data skor dengan metode penskoran yang berbeda yaitu data skor menggunakan metode penskoran *number-right score*, data skor menggunakan metode penskoran *punishment score* dan data skor menggunakan metode penskoran *reward score*.
3. Kelompok data skor tersebut menjadi populasi data skor penelitian
4. Lakukan analisis faktor untuk mendapatkan jumlah data skor minimum yang memadai untuk melakukan penelitian.
5. Setelah jumlah sampel minimum ditentukan, lakukan pengambilan data skor secara acak dengan pengembalian.
6. Pengambilan sampel acak skor responden dengan pengembalian sebanyak 10 kali jumlah variabel pada analisis faktor yang telah dilakukan sebelumnya.
7. Dari pengambilan acak tersebut hitung nilai reliabilitas dari masing-masing metode penskoran.
8. Didapat nilai reliabilitas dari tiga metode penskoran yang berbeda dari pengambilan sampel acak dengan pengembalian.
9. Lakukan langkah (6.) dan (7.) hingga didapat 30 nilai reliabilitas dari tiga metode penskoran yang berbeda
10. Data 30 buah koefisien reliabilitas dari masing-masing kelompok merupakan data sampel untuk pengujian hipotesis.

E. Teknik Pengolahan Data

1. Analisis Faktor

Maria Agustina Amelia, 2014

KETERBANDINGAN RELIABILITAS TES HASIL BELAJAR MATEMATIKA BERDASAR METODE PENSKORAN NUMBER-RIGHT SCORE DAN METODE PENSKORAN CORRECTION FOR GUESSING

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

Sebelum dilakukan uji asumsi, dilakukan analisis tentang jumlah sampel responden minimum yang memadai untuk melakukan penelitian dan mengetahui berapa jumlah variabel yang terdeteksi menggunakan analisis faktor. Banyaknya variabel digunakan untuk menentukan berapa banyak pengambilan sampel acak dengan pengembalian.

Analisis faktor dilakukan dengan Keiser-Meyer-Olkin (KMO) terlebih dahulu. Jika nilai KMO $> 0,5$ maka analisis lanjutan dapat dilakukan. Analisis tiap butir dilakukan dengan melihat matriks korelasi anti image. Butir-butir dengan korelasi anti image yang bernilai di atas 0,5 dianggap baik dan dapat digunakan dalam pengujian hipotesis.

“Dalam analisis faktor dikenal dua macam prosedur yang dilandasi oleh dasar pikiran yang agak berbeda yaitu *eksplanatory factor analysis* dan *confirmatory factor analysis*” (Azwar, 2013: 122). Analisis faktor yang dilakukan pada penelitian ini merupakan prosedur *eksplanatory factor analysis* karena “membantu peneliti untuk mengenali dan mengidentifikasi berbagai faktor yang membentuk suatu konstruk dengan cara menemukan varians skor terbesar dengan jumlah faktor paling sedikit” (Azwar, 2013: 122). Selain itu dikatakan prosedur *eksplanatory factor analysis* karena tidak bermaksud menguji hipotesis mengenai banyaknya variabel yang terbentuk dari data skor tes hasil belajar. Menurut Crocker dan Algina (1986: 304) “suatu analisis faktor dikatakan eksplanatori jika peneliti tidak memiliki hipotesis mengenai jumlah faktor yang terukur pada suatu tes”.

Tujuan dilakukan analisis faktor adalah mendapatkan jumlah sampel minimum yang memenuhi syarat KMO dan dengan butir-butir yang telah mencakup seluruh variabel yang ingin diteliti. Analisis faktor juga dilakukan untuk mendapatkan banyaknya pengambilan sampel acak dengan pengembalian yang digunakan untuk mendapatkan nilai reliabilitas skor hasil belajar sebagai data sampel. Sejalan dengan pendapat Roscoe (dalam Taniredja, 2012: 38) “Bila dalam penelitian akan melakukan analisis dengan multivariate, maka jumlah anggota sampel minimal 10 kali jumlah variabel yang diteliti”.

2. Uji persyaratan

Uji normalitas

Uji normalitas data digunakan untuk melihat apakah distribusi data sampel penelitian berasal dari populasi berdistribusi normal. Data yang berdistribusi normal merupakan salah satu syarat sebelum menggunakan uji parametrik. Uji normalitas dilakukan menggunakan uji Saphiro-Wilk. Uji Saphiro-Wilk dipilih karena “Uji Saphiro-Wilk merupakan uji yang paling kuat meskipun tidak cukup efektif untuk ukuran sampel kecil” (Razali dan Bee, 2011).

Langkah-langkah uji Saphiro-Wilk adalah

Hipotesis uji

H_0 : data sampel berdistribusi normal

H_1 : data sampel berdistribusi tidak normal

Kriteria uji

H_0 ditolak jika $p < \alpha = 0,05$

H_0 diterima jika $p > \alpha = 0,05$

3. Uji hipotesis

Uji hipotesis dilakukan dengan menggunakan uji perbedaan rata-rata untuk sampel berpasangan untuk tiga sampel jika data telah teruji normalitasnya. Karena uji perbedaan rata-rata untuk tiga sampel berpasangan tidak ada untuk uji parametrik maka digunakan uji non parametrik Friedman test untuk menguji perbedaan rata-rata untuk tiga sampel berpasangan.

Hipotesis uji

H_0 : $\mu_{NR} = \mu_P = \mu_R$ (Tidak ada perbedaan rata-rata nilai reliabilitas skor hasil belajar dengan metode penskoran *number-right score*, *punishment score* dan *reward score*)

H_1 : minimal ada satu pasang nilai reliabilitas skor hasil belajar dengan metode penskoran *number-right score*, *punishment score* dan *reward score* yang berbeda

Kriteria uji

H_0 ditolak jika $p < \alpha = 0,05$

H_0 diterima jika $p > \alpha = 0,05$

Jika pada uji hipotesis disimpulkan ada perbedaan rata-rata antara reliabilitas dengan metode penskoran yang berbeda, maka perlu ada uji *Posthoc* untuk mencari tahu pasangan yang memiliki perbedaan yang signifikan.

Uji *Posthoc*

Uji *Posthoc* digunakan untuk melakukan pengujian pasca uji hipotesis. Uji hipotesis dilakukan untuk menguji perbedaan rata-rata untuk tiga sampel berpasangan. Karenanya belum diketahui seberapa besar perbedaan yang terjadi dan sampel mana yang dikatakan berbeda atau sampel mana yang dikatakan tidak berbeda. Uji *posthoc* dilakukan dengan melakukan uji perbedaan rata-rata untuk dua sampel berpasangan secara berulang. Uji *posthoc* yang sesuai ada dua jenis tergantung pada dipenuhinya asumsi data sebagai berikut:

a. Data berasal dari populasi berdistribusi normal

Jika kedua data berdistribusi normal maka uji *Posthoc* menggunakan uji parametrik yaitu uji-t untuk perbedaan rata-rata dua sampel berpasangan. Uji-t tersebut dilakukan secara berulang sebanyak tiga kali. Uji perbedaan yang dilakukan untuk menguji antara :

- 1) Apakah ada perbedaan reliabilitas hasil belajar antara yang dinilai menggunakan metode penskoran *number-right score* dengan metode penskoran *punishment score*?
- 2) Apakah ada perbedaan reliabilitas hasil belajar antara yang dinilai menggunakan metode penskoran *number-right score* dengan metode penskoran *reward score*?
- 3) Apakah ada perbedaan reliabilitas hasil belajar antara yang dinilai menggunakan metode penskoran *punishment score* dengan metode penskoran *reward score*?

Maria Agustina Amelia, 2014

KETERBANDINGAN RELIABILITAS TES HASIL BELAJAR MATEMATIKA BERDASAR METODE PENSKORAN NUMBER-RIGHT SCORE DAN METODE PENSKORAN CORRECTION FOR GUESSING

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

Karena uji *Posthoc* dilakukan menggunakan uji-t untuk perbedaan rata-rata dua sampel berpasangan yang diulang sebanyak tiga kali akan terjadi kenaikan kesalahan tipe I. Untuk mengatasi kenaikan kesalahan tipe I karena perulangan yang dilakukan, maka digunakan prosedur Bonferroni yang akan mengubah nilai α yang digunakan.

Perulangan dilakukan sebanyak tiga kali. $K=3$, dengan $\alpha = 0,05$.

$$\alpha_B = \frac{0,05}{3} = 0,0167$$

Hipotesis uji

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$

Kriteria uji

H_0 ditolak jika $p < \alpha = 0,05$

H_0 diterima jika $p > \alpha = 0,05$

b. Data berasal dari populasi berdistribusi tidak normal

Jika salah satu dari kedua data yang dibandingkan berdistribusi tidak normal maka uji *Posthoc* menggunakan uji non-parametrik yaitu uji Wilcoxon untuk perbedaan rata-rata dua sampel berpasangan. Uji Wilcoxon tersebut dilakukan secara berulang sebanyak tiga kali. Uji perbedaan yang dilakukan untuk menguji antara :

- 1) Apakah ada perbedaan reliabilitas hasil belajar antara yang dinilai menggunakan metode penskoran *number-right score* dengan metode penskoran *punishment score*?
- 2) Apakah ada perbedaan reliabilitas hasil belajar antara yang dinilai menggunakan metode penskoran *number-right score* dengan metode penskoran *reward score*?

Maria Agustina Amelia, 2014

KETERBANDINGAN RELIABILITAS TES HASIL BELAJAR MATEMATIKA BERDASAR METODE PENSKORAN NUMBER-RIGHT SCORE DAN METODE PENSKORAN CORRECTION FOR GUESSING

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

- 3) Apakah ada perbedaan reliabilitas hasil belajar antara yang dinilai menggunakan metode penskoran *punishment score* dengan metode penskoran *reward score*?

Uji *Posthoc* dilakukan untuk mengetahui pasangan reliabilitas yang memiliki perbedaan yang signifikan. Karena uji *Posthoc* dilakukan menggunakan uji wilcoxon untuk perbedaan rata-rata dua sampel berpasangan yang diulang sebanyak tiga kali akan terjadi kenaikan kesalahan tipe I. Untuk mengatasi kenaikan kesalahan tipe I karena perulangan yang dilakukan, maka digunakan prosedur Bonnferoni yang akan mengubah nilai α yang digunakan.

Perulangan dilakukan sebanyak tiga kali. $K=3$, dengan $\alpha = 0,05$.

$$\alpha_B = \frac{0,05}{3} = 0,0167$$

Hipotesis uji

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$

Kriteria uji

H_0 ditolak jika $p < \alpha = 0,05$

H_0 diterima jika $p > \alpha = 0,05$