

## **ANALISIS BUTIR SOAL UJIAN TENGAH SEMESTER MATA KULIAH ANALISIS DATA STATISTIK**

**Yumi Sarassanti<sup>1</sup>, Ali Hasmy<sup>2</sup>, Desty Septianawati<sup>3</sup>, Nopita Sari<sup>4</sup>**

<sup>1,2,3</sup>Tadris Matematika, FTIK, IAIN Pontianak

<sup>4</sup>Pendidikan Agama Islam, FTIK, IAIN Pontianak

<sup>1</sup>Alamat e-mail yumisarassanti@yahoo.co.id

### **Abstrak**

Tujuan penelitian ini adalah untuk mendeskripsikan karakteristik butir soal Ujian Tengah Semester (UTS) mata kuliah Analisis Data Statistik. Metode penelitian yang digunakan dalam penelitian ini adalah penelitian kuantitatif deskriptif. Teknik pengumpulan data yang dilakukan dengan tes dengan bentuk soal *essay*. Sumber data pada penelitian adalah 21 lembar jawaban UTS mahasiswa prodi Tadris Matematika IAIN Pontianak. Hasil penelitian menunjukkan beberapa hal sebagai berikut: 1) Validitas isi dari 2 validator dengan 5 kategori untuk 12 soal seluruhnya mendapat skor 1 (valid); 2) Tingkat kesulitan butir untuk 12 butir soal, terdapat 3 butir kategori mudah (30%), 8 butir kategori sedang (66,7%), dan 1 butir kategori sulit/sukar (3,3%); 3) Untuk Tingkat daya pembeda, terdapat 10 butir kategori daya pembedanya sangat tinggi (83,3%), 2 butir kategori daya pembedanya sedang (16,7%); 4) Reliabilitas antarrater didapat koefisien korelasi sebesar 0,776 termasuk kategori Reliabel; 5) Sedangkan untuk validitas konstruks menggunakan factor Analysis seluruh butir memiliki koefisien > 0,5 (valid).

**Kata Kunci:** Butir Soal, Validitas, Realibilitas, Tingkat Kesukaran, Daya Pembeda.

### **Abstract**

*The purpose of this study was to describe the characteristics of the Mid Semester Examination (UTS) items in the Statistical Data Analysis course. The research method used in this research is descriptive quantitative research. Data collection techniques were carried out by tests in the form of essay questions. The source of the data in this study were 21 sheets of UTS answers from Tadris Mathematics study program students at IAIN Pontianak. The results of the study show the following: 1) The content validity of the 2 validators with 5 categories for 12 questions all scored 1 (valid); 2) Item difficulty level for 12 items, there are 3 items in the easy category (30%), 8 items in the moderate category (66.7%), and 1 item in the difficult/difficult category (3.3%); 3) For the level of discrimination, there are 10 items in the category of very high discrimination (83.3%), 2 items in the category of moderate discrimination (16.7%); 4) Interrater reliability obtained a correlation coefficient of 0.776 including the Reliable category; 5) As for construct validity using factor analysis, all items have a coefficient > 0.5 (valid).*

**Keywords:** Items, Validity, Reliability, Difficulty Level, Discriminating Power .

## **PENDAHULUAN**

Analisis statistik adalah sebuah proses pengolahan data yang meliputi pemeriksaan, pembersihan, transformasi, pemrosesan, dan pemodelan data. Tujuannya adalah untuk menemukan informasi yang berguna dan dapat digunakan sebagai dasar dalam pengambilan kesimpulan atau keputusan. Proses ini menerapkan metode atau teknik statistik sesuai dengan data yang dimiliki. Mengapa analisis data itu penting? Tidak semua orang paham membaca data yang sangat banyak dan masih dalam bentuk mentah. Dengan melakukan analisis data, kumpulan data mentah akan

diolah sedemikian rupa sehingga mudah dimengerti dan dipahami oleh orang lain. Evaluasi pembelajaran adalah suatu proses atau kegiatan yang bersifat sistematis, berkelanjutan, serta menyeluruh dalam pengendalian, penjaminan, dan penetapan kualitas pembelajaran terhadap berbagai komponen pembelajaran yang berdasarkan pertimbangan dan kriteria tertentu (Arifin, 2014). Guru harus melakukan evaluasi terhadap hasil belajar dan menentukan standar kompetensi yang harus dicapai siswa. Dengan melakukan evaluasi terhadap hasil belajar, guru dapat mengetahui instrumen yang digunakan terlalu mudah atau sulit, atau sudah sesuai dengan indikator dari pembelajaran atau tidak, serta pembelajaran (model, pendekatan, strategi, dan metode) yang digunakan dalam mengajar. Jika instrumen yang digunakan terlalu sulit, maka perlu dilakukan perbaikan dengan cara menganalisis setiap butir soal yang digunakan. Analisis butir soal merupakan proses mengkaji terhadap kualitas soal pada setiap butirnya. Menganalisis setiap butir soal sangat penting dilakukan, jangan sampai setiap butir soal terdapat hal-hal yang tidak sesuai dengan tujuan pembelajaran, jika dilihat dari tingkat kesukaran, pola pengecoh, daya pembeda, dan yang lainnya. Banyak guru atau calon guru yang telah memiliki kemampuan yang memadai tentang bagaimana cara merumuskan tujuan pembelajaran, menentukan bahan pelajaran secara terperinci, memilih dan menentukan metodologi mengajar, menyiapkan media dan sarana pembelajaran, tetapi masih belum memiliki kemampuan yang diharapkan dalam bidang evaluasi hasil belajar khususnya dalam menganalisis butir soal (Validitas, Reliabilitas, tingkat kesukaran, pola pengecoh, daya pembeda). Oleh karena itu, kemampuan dan keterampilan melakukan evaluasi merupakan kemampuan profesional yang harus dikuasai oleh setiap guru atau calon guru. Menguasai kemampuan ini tidaklah mudah, tetapi memerlukan latihan dan pengalaman lapangan yang memadai. Kemampuan dalam bidang evaluasi, khususnya evaluasi hasil belajar bukan hanya bermanfaat dalam proses belajar mengajar, tetapi juga bermanfaat dalam rangka penelitian ilmiah, yaitu tentang bagaimana cara membuat alat ukur yang valid dan reliabel. Dalam menyusun alat ukur diperlukan beberapa langkah, yaitu: cara menyusun/menulis alat ukur serta menganalisis alat ukur sehingga menjadi alat ukur yang valid dan reliabel (Mahendra, 2019). Analisis butir soal adalah sebuah proses yang harus ditempuh dan dilaksanakan oleh seorang guru untuk mengetahui sejauh mana kualitas butir soal yang diberikan dan dengan adanya kegiatan analisis ini bisa membantu para guru terkait butir soal yang bagus dan yang layak untuk dipertahankan dan butir soal yang harus dibuang (Farida dan Musyarofah.A. 2021). Oleh karena itu, penelitian ini bertujuan untuk mendeskripsikan karakteristik butir soal Ujian Tengah Semester (UTS) mata kuliah Analisis Data Statistik.

## METODE

Penelitian ini merupakan penelitian kuantitatif deskriptif. Menurut Lufri dan Ardi (2014), penelitian deskriptif merupakan penelitian yang mendeskripsikan peristiwa, gejala, fakta, kejadian yang sudah atau sedang terjadi. Subjek dalam penelitian ini adalah 21 orang mahasiswa Angkatan pertama Tadris Matematika IAIN Pontianak. Teknik pengumpulan data yang digunakan yaitu melalui instrumen tes berupa soal esai, lembar jawaban, dan kunci jawaban. Analisis data penelitian menggunakan Excel dan Spss untuk mengetahui validitas, reliabilitas, tingkat kesukaran, daya beda. Adapun tahapan analisis butir soal sebagai berikut :

Reliabilitas merupakan ukuran yang menyatakan tingkat kekonsistenan suatu soal. Reliabilitas berhubungan dengan apakah suatu tes dapat dipercaya sesuai dengan kriteria yang telah ditetapkan. Berikut adalah kriteria untuk menafsirkan reliabilitas soal pada Tabel 1 (Putri dan Ofianto, 2019).

**Tabel 1. Kriteria Reliabilitas Soal**

Nilai reliabilitas	Interpretasi
0.800-1.000	Sangat Tinggi
0.600-0.799	Tinggi
0.400-0.599	Cukup
0.200-0.399	Rendah
0.000-0.199	Sangat Rendah

Tingkat kesukaran adalah pengukuran seberapa besar kesukaran suatu soal. Salah satu syarat instrumen tes yang baik adalah memiliki tingkat kesukaran yang tidak terlalu sulit dan tidak terlalu mudah. Adapun kriteria penafsiran tingkat kesukaran soal dapat dilihat pada Tabel 2 (Arifin, 2012).

**Tabel 2. Kriteria Tingkat Kesukaran Soal**

Tingkat Kesukaran (%)	Interpretasi
0 – 27	Sukar
28 - 72	Sedang
73 – 100	Mudah

Daya pembeda adalah pengukuran suatu soal dalam membedakan kemampuan peserta didik, antara peserta didik yang memiliki kemampuan tinggi dengan kemampuan rendah. Untuk menginterpretasikan koefisien daya pembeda, dapat digunakan kriteria pada Tabel 3 (Elviana, 2020).

**Tabel 3. Kriteria Tingkat Daya Pembeda Soal**

Tingkat Daya Pembeda (%)	Interpretasi
Negatif – 9	Sangat Buruk (Harus Dibuang)
10 – 19	Buruk (Sebaiknya Dibuang)
20 – 29	Agak Baik (Cukup)
30 – 49	Baik
50 ke Atas	Sangat Baik

## HASIL DAN PEMBAHASAN

Peneliti melakukan uji validitas konstruk, data di lapangan yang telah terkumpul kemudian dianalisis dengan menggunakan *Software Partial Least Square* (PLS) guna menguji valid atau tidaknya setiap butir pernyataan instrumen yang digunakan. Adapun gambar dari analisis *Software Partial Least Square* (PLS) setiap butir dapat dilihat uraian berikut ini:

### Analisis Butir Soal

Variable	Weight	Loading	Location	ResidVar	Redundan
X <sub>1</sub> outward					
X <sub>1.1</sub>	0.151	0.791	.0000	0.366	.0000
X <sub>1.2</sub>	0.147	0.735	.0000	0.451	.0000
X <sub>1.3</sub>	0.141	0.728	.0000	0.461	.0000
X <sub>2</sub> outward					
X <sub>1.4</sub>	0.198	0.788	.0000	0.370	.0000
X <sub>1.5</sub>	0.200	0.802	.0000	0.351	.0000
X <sub>1.6</sub>	0.182	0.753	.0000	0.445	.0000
X <sub>3</sub> outward					
X <sub>1.7</sub>	0.204	0.766	.0000	0.404	.0000
X <sub>1.8</sub>	0.216	0.837	.0000	0.304	.0000
X <sub>1.9</sub>	0.193	0.748	.0000	0.429	.0000
X <sub>1.10</sub>	0.216	0.786	.0000	0.374	.0000
X <sub>1.11</sub>	0.220	0.849	.0000	0.270	.0000
X <sub>1.12</sub>	0.207	0.792	.0000	0.364	.0000

**Gambar 1. Analisis Butir Soal**

Berdasarkan analisis di atas, dapat diketahui lebih jelas terkait kriteria yang dikatakan valid dari pengolahan data analisis butir soal bahwa nilai loading  $\geq 0,5$  maka dapat dikatakan butir soal tersebut valid. Dari gambar di atas untuk analisis butir soal pada materi mean, median, modus yang

terdapat dari butir nomor 1-3, dapat dilihat bahwa nilai loading untuk semua butir tersebut  $\geq 0,5$  maka dapat dikatakan butir soal tersebut valid. Maka kesimpulannya setiap butir soal pada data analisis butir soal pada materi mean, median, modus dikatakan valid.

Dari pengolahan data analisis butir soal pada materi kuartil, desil, dan persentil yang terdapat dari butir nomor 4-6, dapat dilihat bahwa nilai loading untuk semua butir tersebut  $\geq 0,5$  maka dapat dikatakan butir soal tersebut valid. Maka kesimpulannya setiap butir soal pada data analisis butir soal pada materi kuartil, desil, persentil dikatakan valid. Kemudian dari pengolahan data untuk analisis butir soal pada materi rentang, rentang antar kuartil, rentang semi antarkuartil, simpangan rata-rata, simpangan baku, dan varian yang terdapat dari butir nomor 7-12, dapat dilihat bahwa nilai loading untuk semua butir Kompetensi Kepribadian Dosen  $\geq 0,5$  maka dapat dikatakan butir soal tersebut valid. Maka kesimpulannya setiap butir soal pada data analisis butir soal pada materi rentang, rentang antar kuartil, rentang semi antarkuartil, simpangan rata-rata, simpangan baku, dan varian dikatakan valid.

**Tabel 4. Perhitungan reliabel**

Cronbach's Alpha		
Cronbach's Alpha	Based on Standardized Items	N of Items
,776	,881	12

Berdasarkan tabel di atas didapat bahwa Reliabilitas antarrater didapat koefisien korelasi sebesar 0,776 termasuk kategori Reliabeli. Hal ini sesuai dengan pendapat Ratnawulan dan Rusdiana (2014) yang menyatakan bahwa, salah satu syarat tes sebagai salah satu instrumen evaluasi adalah memiliki reliabilitas yang tinggi. Tes yang memiliki reliabilitas tinggi akan memberikan hasil yang tetap dan tidak berubah-ubah.

Berdasarkan hasil analisis tingkat kesukaran butir soal, bahwa tingkat kesulitan butir untuk 12 butir soal, terdapat 3 butir kategori mudah (30%), 8 butir kategori sedang (66,7%), dan 1 butir kategori sulit/sukar (3,3%); Hal ini sesuai dengan pendapat Arifin (2012) bahwa, suatu soal tes hendaknya tidak terlalu sukar dan tidak pula terlalu mudah. Soal yang memiliki tingkat kesukaran seimbang (proporsional), maka dapat dikatakan bahwa soal tersebut baik. Tingkat Daya Pembeda Soal Hasil analisis untuk Tingkat daya pembeda, terdapat 10 butir kategori daya pembedanya sangat tinggi (83,3%) dan 2 butir kategori daya pembedanya sedang (16,7%), Hal ini sesuai dengan pendapat Arifin (2014) bahwa, semakin tinggi koefisien daya pembeda suatu butir soal, maka

---

---

semakin mampu butir soal tersebut membedakan antara peserta didik yang mampu menguasai dengan yang kurang mampu menguasai kompetensi. Dijelaskan oleh Alpusari (2014), butir soal yang tergolong kategori dipakai, maka butir soal tersebut langsung digunakan dan dimasukkan dalam bank soal yang ada pada guru, dan soal tersebut dapat digunakan lagi dalam tes hasil belajar yang akan datang. Butir soal yang termasuk ke dalam kategori dipakai diperbaiki, maka soal tersebut diteliti dan diperbaiki tata bahasanya, apakah kalimat soal yang kurang jelas atau sulit dipahami oleh peserta didik. Sedangkan butir soal yang tidak dipakai dapat dipakai langsung dibuang dan tidak digunakan untuk tes hasil belajar berikutnya. Sejalan dengan temuan (Santoso, dkk. 2019) penelitian ini memberikan beberapa implikasi bagi praktik maupun penelitian selanjutnya. Pengembang tes hendaknya benar-benar memperhatikan kaidah dalam penyusunan instrumen tes, sehingga dapat dihasilkan butir-butir tes yang berkualitas menguasai kaidah-kaidah dasar dalam menyusun butir tes. Penelitian-penelitian ke depannya hendaknya juga menganalisis kualitas instrumen tes yang dikembangkan oleh lembaga/lembaga lainnya, terutama untuk tes yang dikembangkan secara luas. Penelitian terkait kualitas tes tidak terbatas hanya pada instrumen berbentuk objektif, tetapi juga dapat dilakukan pada instrumen tes berbentuk uraian maupun bentuk tes lainnya.

## SIMPULAN

Berdasarkan hasil analisis butir soal ujian akhir semester analisis data statistik pada semester dapat disimpulkan bahwa kualitas soal tergolong baik. Hal ini ditunjukkan dengan hasil analisis butir soal dari 12 soal yang diujikan kepada 21 mahasiswa. kelas XI diketahui bahwa, 1) Validitas isi dari 2 validator dengan 5 kategori untuk 12 soal seluruhnya mendapat skor 1 (valid); 2) Tingkat kesulitan butir untuk 12 butir soal, terdapat 3 butir kategori mudah (30%), 8 butir kategori sedang (66,7%), dan 1 butir kategori sulit/sukar (3,3%); 3) Untuk Tingkat daya pembeda, terdapat 10 butir kategori daya pembedanya sangat tinggi (83,3%), 2 butir kategori daya pembedanya sedang (16,7%); 4) Reliabilitas antarrater didapat koefisien korelasi sebesar 0,776 termasuk kategori Reliabel; 5) Sedangkan untuk validitas konstruks menggunakan factor Analysis seluruh butir memiliki koefisien > 0,5 (valid).

## DAFTAR PUSTAKA

Alpusari, M. (2014). Analisis Butir Soal Konsep Dasar IPA 1 Melalui Penggunaan Program Komputer Anates Versi 4.0 For Windows. Primary: *Jurnal Pendidikan Guru Sekolah Dasar*, 3(2), 106-115.

- Arifin, Z. (2012). *Evaluasi Pembelajaran Penulis*. Jakarta: Direktorat Jenderal Pendidikan Islam Kementerian Agama.
- Arifin, Z. (2014). *Penelitian Pendidikan Metode dan Paradigma Baru*. Bandung: PT Remaja Rosdakarya.
- Elviana. (2020). Analisis Butir Soal Evaluasi Pembelajaran Pendidikan Agama Islam Menggunakan Program Anates. *Jurnal Mudarrisuna*. 10(2), 58-74.
- Farida dan Musyarofah.A. (2021). Validitas dan Reliabilitas dalam Analisis Butir Soal. Al-Mu'arrib. *Jurnal Pendidikan Bahasa Arab*. 1(1), 34-44.
- Lufri, L., dan Ardi, A. (2014). *Metodologi Penelitian*. Padang: UNP Press.
- Mahendra, I Wayan Eka. 2019. *Analisis Butir Soal*. Bali : FPMIPA IKIP PGRI Bali
- Putri, R.H., dan Ofianto. (2019). Efektivitas Analisis Butir Menggunakan Anajohn, Anates, dan Iteman Studi Soal USBN Pelajaran Sejarah Kota Padang. *Kronologi : Jurnal Mahasiswa Ilmu Sejarah dan Pendidikan*, 1(2), 1-11.
- Santoso. A. dkk. (2019). Kualitas Butir Bank Soal Statistika (Studi Kasus: Instrumen Ujian Akhir Mata Kuliah Statistika Universitas Terbuka). *Jurnal Riset Pendidikan Matematika*. 6 (2), 165-176 .