

Two-Sample t-Test



OLEH:

CHANDRA NOVTIAR, M.SI

PROGRAM STUDI PENDIDIKAN MATEMATIKA FAKULTAS PENDIDIKAN

MATEMATIKA DAN SAINS

INSTITUT KEGURUAN DAN ILMU PENDIDIKAN

(IKIP) SILIWANGI

2019

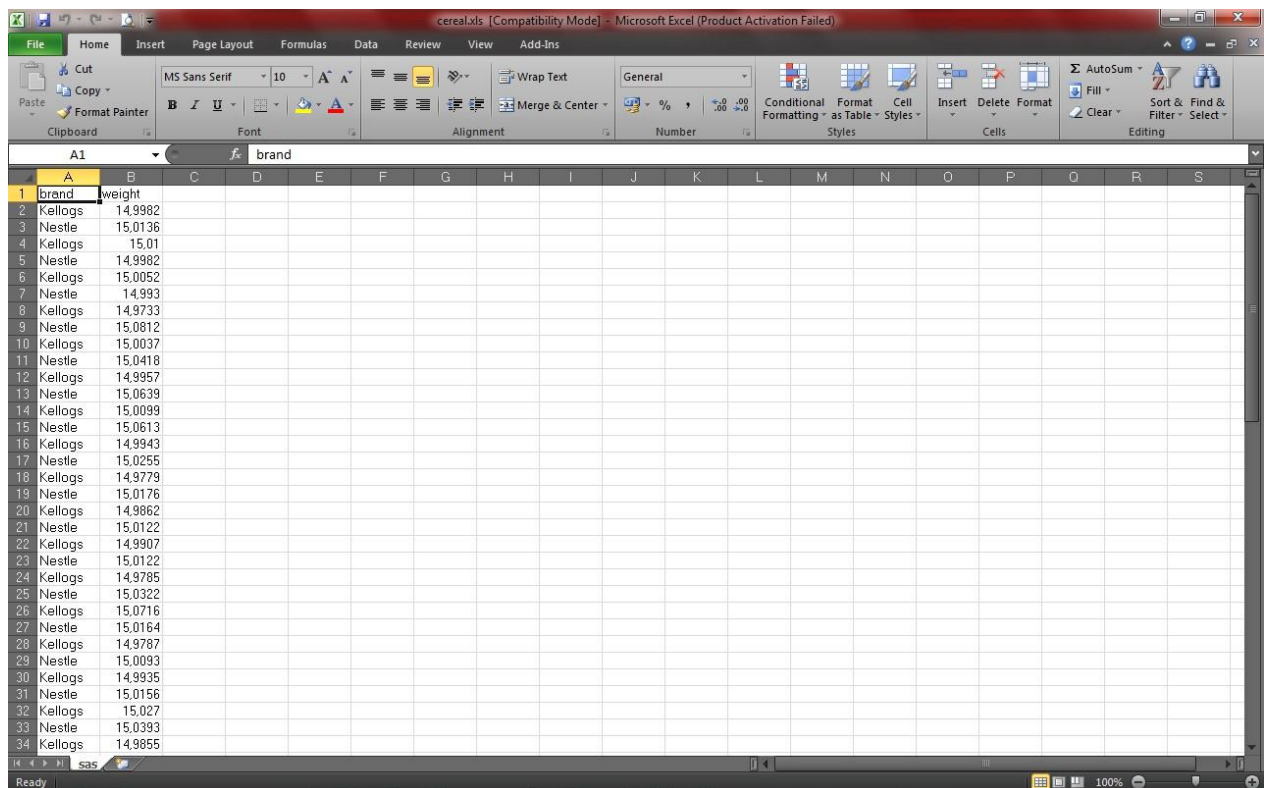
1Masalah

Pabrik yang sama memproduksi **2** jenis cereal yaitu Kellog dan Nestle. Pihak pabrik ingin meyakinkan bahwa proses pembuatan cereal yang berbeda ini memasukkan kuantitas (berat) yang sama dalam setiap box cerealnya. Setiap merk tersebut diharuskan memiliki berat **15** ons/box. Sampel acak dari setiap merk dipilih dan massa dari sereal tersebut dicatat.

- Data disimpan dalam data set dengan nama **cereal**.
- Variabel dalam data set adalah :
 - brand** dua grup, Kellog dan Nestle, yang berkorespondensi dengan **2**(dua) merk
 - weight** berat sereal dalam ons

2Data

Data berat sereal dalam file "cereal.xls" sebagai berikut :



brand	weight
Kellogs	14.9982
Nestle	15.0136
Kellogs	15.01
Nestle	14.9982
Kellogs	15.0052
Nestle	14.993
Kellogs	14.9733
Nestle	15.0812
Kellogs	15.0037
Nestle	15.0418
Kellogs	14.9957
Nestle	15.0639
Kellogs	15.0099
Nestle	15.0613
Kellogs	14.9943
Nestle	15.0255
Kellogs	14.9779
Nestle	15.0176
Kellogs	14.9862
Nestle	15.0122
Kellogs	14.9907
Nestle	15.0122
Kellogs	14.9785
Nestle	15.0322
Kellogs	15.0716
Nestle	15.0164
Kellogs	14.9787
Nestle	15.0093
Kellogs	14.9935
Nestle	15.0156
Kellogs	15.027
Nestle	15.0393
Kellogs	14.9855

3Pengolahan Data, Analisa dan Pembahasan

file "cereal.xls" diimport ke dalam program SAS kemudian disimpan dalam library data dalam data set cereal. Sebagian output data ditunjukkan oleh gambar :

Listing of Cereal Data

brand	weight
Kellogs	14.9982
Nestle	15.0136
Kellogs	15.0100
Nestle	14.9982
Kellogs	15.0052
Nestle	14.9930
Kellogs	14.9733
Nestle	15.0812
Kellogs	15.0037
Nestle	15.0418
Kellogs	14.9957
Nestle	15.0639
Kellogs	15.0099
Nestle	15.0613
Kellogs	14.9943
Nestle	15.0255
Kellogs	14.9779
Nestle	15.0176
Kellogs	14.9862
Nestle	15.0122
Kellogs	14.9907
Nestle	15.0122
Kellogs	14.9785
Nestle	15.0322
Kellogs	15.0716
Nestle	15.0164

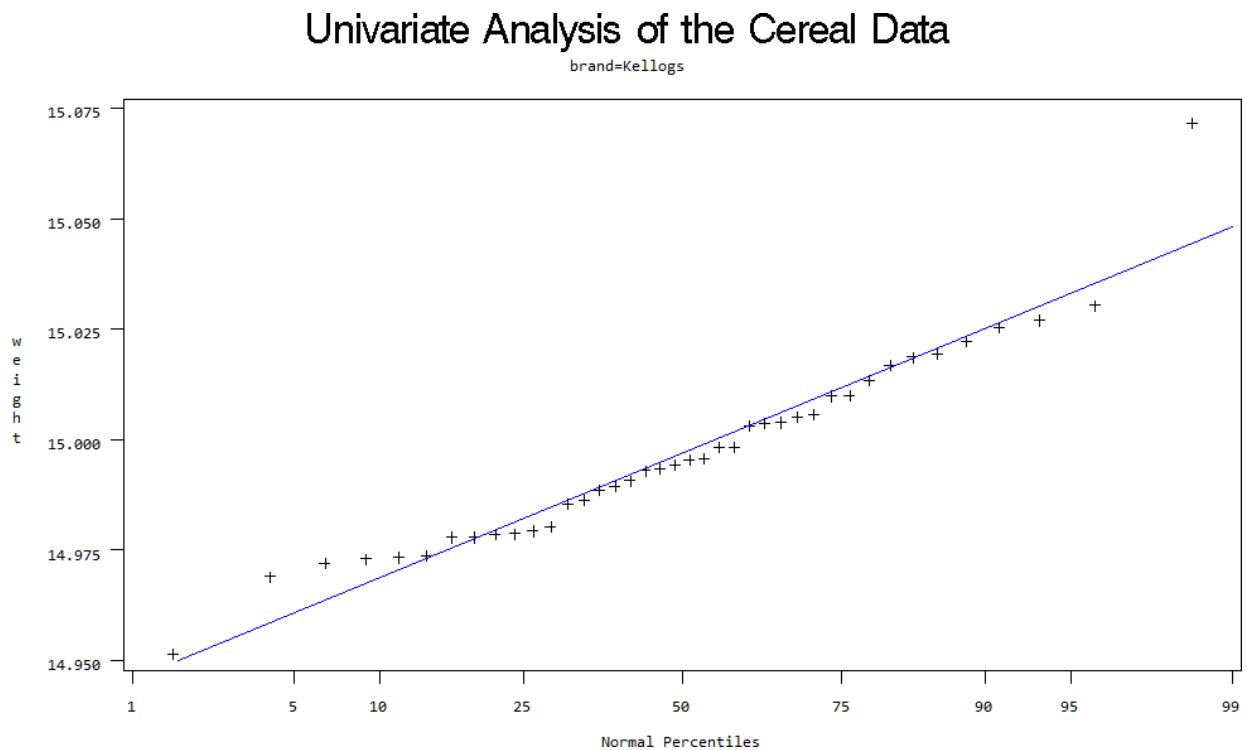
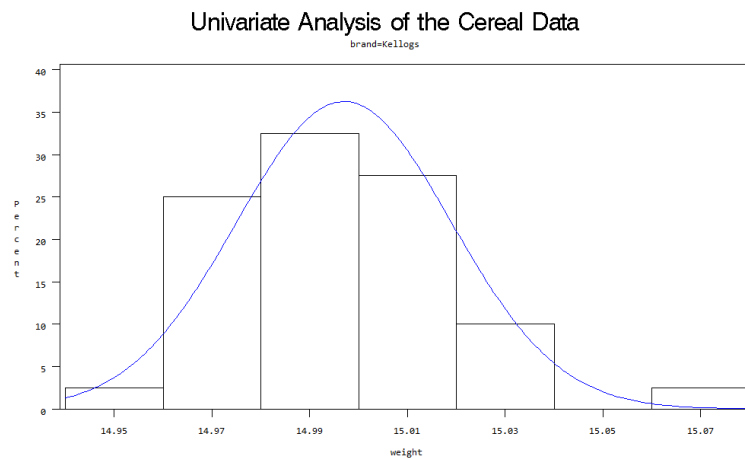
Pada tugas kali ini, akan dibandingkan bahwa rata-rata berat cereal dalam setiap box antara merk Kellog dan Nestle sama. Untuk membandingkannya digunakan statistik uji t (t -test). Untuk membandingkan rata-rata berat sereal Kellog dan Nestle, diasumsikan bahwa ;

- Observasi saling bebas;
- Data dari setiap grup berdistribusi normal; dan
- Variansi setiap grup sama.

3.1 Normality

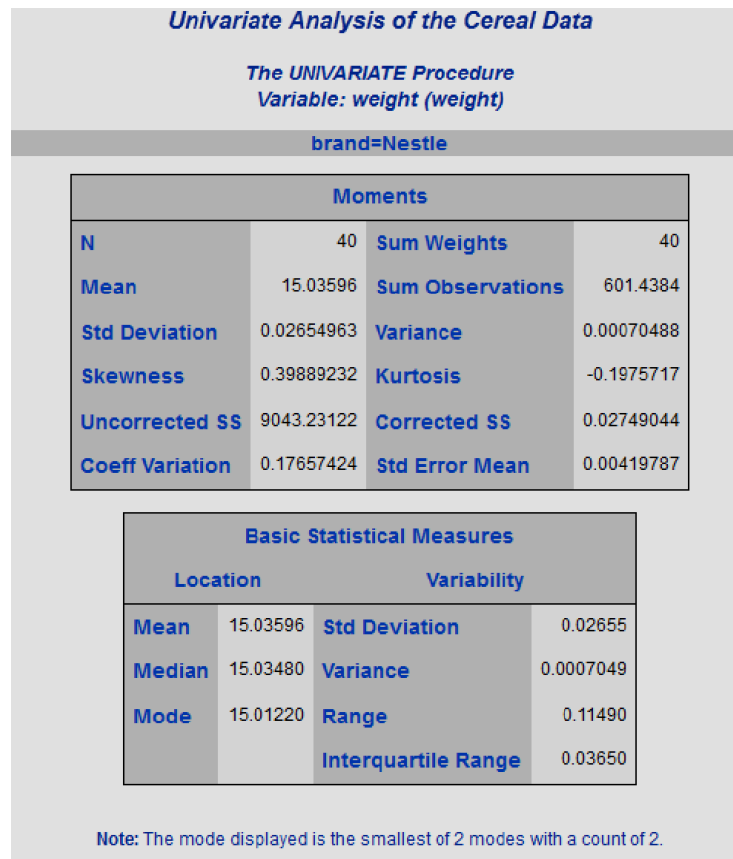
3.1.1 Kellogs

Univariate Analysis of the Cereal Data			
The UNIVARIATE Procedure			
Variable: weight (weight)			
brand=Kellogs			
Moments			
N	40	Sum Weights	40
Mean	14.9970125	Sum Observations	599.8805
Std Deviation	0.02201048	Variance	0.00048446
Skewness	0.87481049	Kurtosis	2.07993397
Uncorrected SS	8996.43425	Corrected SS	0.01889398
Coeff Variation	0.14676575	Std Error Mean	0.00348016
Basic Statistical Measures			
Location		Variability	
Mean	14.99701	Std Deviation	0.02201
Median	14.99490	Variance	0.0004845
Mode	14.97790	Range	0.12010
		Interquartile Range	0.03095
Note: The mode displayed is the smallest of 2 modes with a count of 2.			

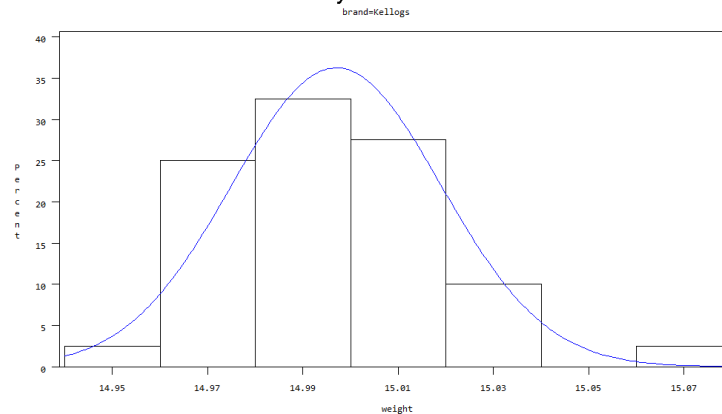


Histogram dan normal probability plots menunjukkan satu nilai ekstrim. Oleh karena itu, data Kellogs terlihat simetris. Hal ini menunjukkan bahwa tidak ada pola untuk data yang merefleksikan skewness dan kurtosis.

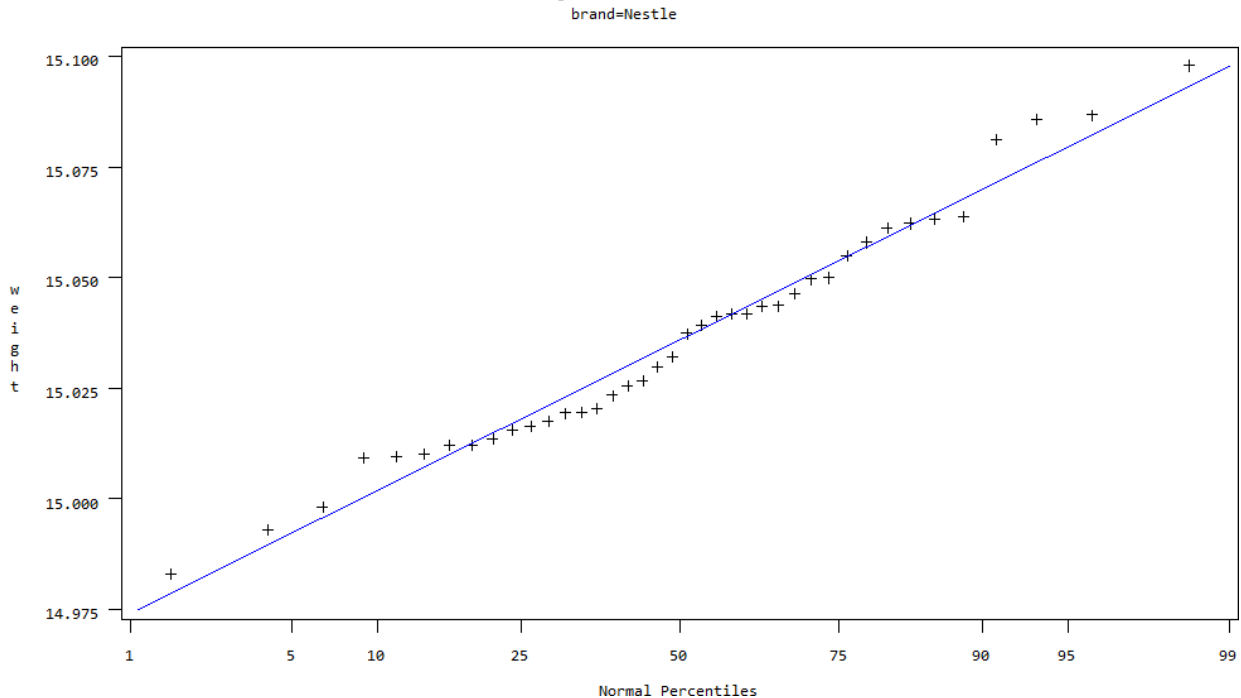
3.1.2 Nestle



Univariate Analysis of the Cereal Data



Univariate Analysis of the Cereal Data



Histogram memperlihatkan bahwa data "fairly symmetric". Dari histogram juga "normal probability plot" juga terlihat bahwa tidak terdapat nilai ekstrim. "Normal probability plot" menunjukkan bahwa "tidak ada" data yang benar-benar meninggalkan garis normal ("tidak ada" : masih dianggap berada pada garis normal).

Karena kedua merk memiliki berat yang berdistribusi normal, maka asumsi untuk F-test untuk variansi sama terbukti. Asumsi untuk t-test yang memperhatikan normalitas dari distribusi rata-rata sampel dipenuhi. Kita harus menggunakan teorema limit pusat untuk memvalidasi bahwa asumsi untuk t-test karena kedua merk memiliki 40 observasi.

3.1.3 UJI NORMALITY KOLMOGOROF-SMIRNOF

Uji normalitas data dilakukan untuk mengetahui distribusi data dari suatu variabel yang akan digunakan dalam penelitian. Data yang baik dan layak adalah data yang memiliki distribusi normal. Ada bermacam-macam cara untuk mengetahui normalitas distribusi data, salah satunya menggunakan uji Kolmogorov-Smirnov. Hipotesis yang diajukan untuk tugas kali ini adalah sebagai berikut:

H_0 : Data Berat sereal Kellogs berdistribusi normal.

H_1 : Data Berat sereal Kellogs tidak berdistribusi normal.

dan

H_0 : Data Berat sereal Nestle berdistribusi normal.

H_1 : Data Berat sereal Nestle tidak berdistribusi normal.

Misalkan dipilih taraf signifikansi $\alpha = 0.05$. Digunakan statistik uji deviasi maksimum $D = |S(x) - F(x)|$ dengan $F(x)$ ialah fungsi distribusi frekuensi kumulatif populasi pengamatan dan $S(x)$ ialah fungsi distribusi frekuensi kumulatif sampel. Pengambilan keputusan:

Jika $p\text{-value}(P > D) > 0.05$ maka H_0 diterima.

Jika $p\text{-value}(P > D) < 0.05$ maka H_0 ditolak.

Dilakukan pengujian terhadap rata-rata berat sereal Kellog dan Nestle. Dengan menggunakan SAS, diperoleh :

Univariate Analysis of the Cereal Data				
The UNIVARIATE Procedure				
Fitted Distribution for weight				
brand=Kellogs				
Goodness-of-Fit Tests for Normal Distribution				
Test	Statistic		p Value	
Kolmogorov-Smirnov	D	0.07848687	Pr > D	>0.150
Cramer-von Mises	W-Sq	0.04993559	Pr > W-Sq	>0.250
Anderson-Darling	A-Sq	0.41433821	Pr > A-Sq	>0.250

Univariate Analysis of the Cereal Data				
The UNIVARIATE Procedure				
Fitted Distribution for weight				
brand=Nestle				
Goodness-of-Fit Tests for Normal Distribution				
Test	Statistic		p Value	
Kolmogorov-Smirnov	D	0.09608648	Pr > D	>0.150
Cramer-von Mises	W-Sq	0.05930447	Pr > W-Sq	>0.250
Anderson-Darling	A-Sq	0.38776343	Pr > A-Sq	>0.250

Berdasarkan uji normalitas ini (lihat gambar), rata-rata berat sereal Kellog dan rata-rata berat sereal Nestle $p\text{-value}(P > D) > \alpha = 0.05$, sehingga H_0 diterima. Jadi, dengan tingkat

signifikansi $\alpha = 0.05$, data Berat sereal Kellog dan data berat sereal Nestle berdistribusi normal.

4 Uji Hipotesis Dua Populasi

Pada bagian ini akan dilakukan uji Hipotesis untuk memperlihatkan bahwa rata-rata berat sereal dalam box Kellog maupun box Nestle sama yaitu **15** ons. Terlebih dahulu akan dilihat variansi dari kedua grup ini sama atau tidak.

4.1 Variansi Dua Populasi

Misalkan :

- σ_k^2 merupakan variansi berat sereal Kellog;
- σ_n^2 merupakan variansi berat sereal Nestle;
- s_k^2 merupakan variansi sampel berat sereal Kellog ; dan
- s_n^2 merupakan variansi sampel berat sereal Nestle;

Didefinisikan :

$$H_0 : \sigma_k^2 = \sigma_n^2$$

$$H_1 : \sigma_k^2 \neq \sigma_n^2$$

dengan tingkat signifikansi $\alpha = 0.05$. Gunakan statistik uji-F, yaitu $F = \frac{\max(s_k^2, s_n^2)}{\min(s_k^2, s_n^2)}$

Testing the Equality of Means for Two Cereal Brands											
The TTEST Procedure											
		Statistics									
Variable	brand	N	Lower CL Mean	Upper CL Mean	Lower CL Std Dev	Upper CL Std Dev	Std Err	Minimum	Maximum		
weight	Kellogs	40	14.99	14.997	0.018	0.022	0.0283	0.0035	14.952	15.072	
weight	Nestle	40	15.027	15.036	0.0217	0.0265	0.0341	0.0042	14.983	15.098	
weight	Diff (1-2)		-0.05	-0.039	-0.028	0.0211	0.0244	0.0289	0.0055		

T-Tests					
Variable	Method	Variances	DF	t Value	Pr > t
weight	Pooled	Equal	78	-7.14	<.0001
weight	Satterthwaite	Unequal	75.4	-7.14	<.0001

Equality of Variances					
Variable	Method	Num DF	Den DF	F Value	Pr > F
weight	Folded F	39	39	1.45	0.2460

Pengambilan keputusan:

Jika $p\text{-value}(P > F) > 0.05$ maka H_0 tidak ditolak.

Jika $p\text{-value}(P > F) < 0.05$ maka H_0 ditolak.

Perhatikan (2) pada gambar. Diperoleh $F = 1.45$ dan $p\text{-value}(P > F) = 0.2460$. Karena $p\text{-value} > \alpha = 0.05$, maka H_0 **tidak ditolak**. Jadi, dengan tingkat signifikansi $\alpha = 0.05$, **tidak cukup bukti** untuk menunjukkan bahwa **variansi berat dari sereal Kellog tidak sama dengan berat sereal Nestle**.

4.2 Mean Dua Populasi

Misalkan :

- μ_k merupakan rata-rata berat sereal Kellog;
- μ_n merupakan rata-rata berat sereal Nestle;
- $\bar{x}_k$ merupakan rata-rata sampel berat sereal Kellog ; dan
- $\bar{x}_n$ merupakan rata-rata sampel berat sereal Nestle;

Didefinisikan :

$$H_0 : \mu_k = \mu_n$$

$$H_1 : \mu_k \neq \mu_n$$

dengan tingkat signifikansi $\alpha = 0.05$. Pengambilan keputusan:

Jika $p\text{-value}(P > |t|) > 0.05$ maka H_0 tidak ditolak.

Jika $p\text{-value}(P > |t|) < 0.05$ maka H_0 ditolak.

Gunakan statistik uji-t. Perhatikan (3) pada gambar. Karena kita tidak bisa menunjukkan bahwa variansi berat dari sereal Kellog tidak sama dengan berat sereal Nestle, kita dapat menganggap bahwa variansi keduanya **sama**. Lihat (3) dimana variansi dari kedua grup sama, lihat bulatan merah. Berdasarkan gambar, $t = -7.14$ dan $p\text{-value}(P > |t|) < 0.0001$, sehingga H_0 **ditolak**. Jadi, dengan tingkat signifikansi $\alpha = 0.05$, terdapat perbedaan dalam rata-rata berat sereal antara merk Kellog dan merk Nestle.

Kembali ke tabel statistik (perhatikan (1)) pada gambar. Karena selang kepercayaan untuk rata-rata **(-0.05, -0.028)** yang tidak memuat **0**, maka dapat disimpulkan bahwa terdapat perbedaan yang signifikan antara rata-rata berat kedua sereal tersebut.

5Kesimpulan

Pabrik tidak bisa meyakinkan bahwa kedua produknya yaitu Kellog dan Nestle memiliki berat yang sama yaitu **15** ons dalam setiap boxnya.

6Lampiran

```
libname data'C:\Users\win 7\Desktop\SAS Academy_Statistical Analysis\PR2';  
ods html  
body='cereal.html';  
proc print data=data.cereal label noobs;  
title 'Listing of Cereal Data';  
run;
```

```
/*proc sort*/  
proc sort data=data.cereal out=data.sorted_cereal;  
by brand;  
run;  
ods html  
body='sorted_cereal.html';  
proc print data=data.sorted_cereal label noobs;  
title 'Listing of Sorted Cereal Data';  
run;
```

```
/*Univariate Analysis of Cereal Data*/  
ods html  
body='Univariate_Analysis.html';  
ods select  
Moments  
BasicMeasures;  
proc univariate data=data.sorted_cereal;  
by brand;  
var weight;  
histogram weight / normal;  
probplot weight / normal (mu=est sigma=est  
color=blue w=1);  
title 'Univariate Analysis of the Cereal Data';  
run;
```

```
/*Two Samples t-test*/  
ods html  
body='ttest.html';  
proc ttest data=data.cereal;  
class brand;  
var weight;  
title 'Testing the Equality of Means for Two Cereal'  
      ' Brands';  
run;
```