Mini Statistical Project – Is rice or noodles cheaper (in Bedok)?

Done by: Pang Jin Jia

Introduction


THE STRAITS TIMES SINGAPORE

Single elderly Singaporeans need $1,379 a month to meet basic living standard: Study


THE STRAITS TIMES SINGAPORE

Household incomes, living standards have gone up in five years: Govt survey

23 May 2019: https://www.straitstimes.com/singapore/study-finds-1379-a-month-needed-to-meet-basic-living-standard-for-single-elderly

1 Aug 2019: https://www.straitstimes.com/singapore/household-incomes-living-standards-have-gone-up-in-five-years-govt-survey

Research Question

Is the average cost of a typical Chinese rice dish less than Chinese noodles in Bedok?

The reason behind asking this research question is to find out if it is cheaper to live off rice or noodles in our everyday meals, for anyone who is on a budget for food. With the assumption that household incomes are rising and the population as a whole should be able to keep up with living standards, this assumption does not apply to elderly who are on the brink of retirement or who have already retired (just like my parents). So just how much allowance do I have to give my parents (assuming that they have no savings at all) for basic survival in terms of food? Bedok town, being a town with a significant proportion of elderly, will be a good place to conduct this study.

Data Collection

For clarity, roasted chicken, white chicken, char siew, chicken cutlet, duck rice cost will be recorded for this study. Fishball, wanton, dumpling, prawn noodles cost will be recorded for this study. Base dish size price will be recorded, not the upsized version. Malay, Thai, Korean, Japanese etc types of food will be excluded from the study, only Chinese dishes will be included.

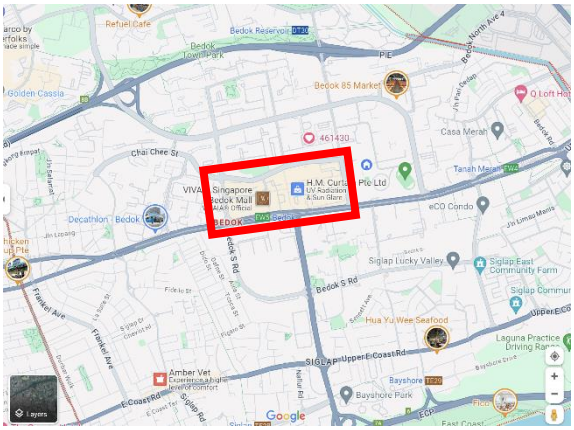
Only places within the vicinity of Bedok town central will be used. Every chicken rice cost, noodles cost found in every hawker center, food court and coffeeshop will be recorded. However, no more than 2 of the least expensive dish per stall will be recorded to ensure spread of data across as many stalls as possible. Expensive shopping mall restaurants rice and noodles prices are omitted. Sample size of both noodles and rice populations will be limited to the number of stalls that can be found within the vicinity that fits the criteria, with at least 20 of each type.

Types of rice and noodles will be recorded. Name of the food court, hawker center or coffeeshop will also be recorded for authenticity check.

Populations: All common types of Chinese rice and noodle dishes in Bedok central

Samples: Common types of Chinese rice and noodle dishes in Bedok central

Sampling method: Cluster sampling (geographically by postal code)



https://en.wikipedia.org/wiki/Bedok

Bedok town is very big, hence only a cluster near my house (pink heart shape) will be sampled (red rectangular box).

Types of data:

Quantitative (discrete): postal code (discrete), price SGD$ (continuous)

Qualitative (nominal): food, location, stall name

Sample table (rice & noodles will be recorded on separate tables):

Food	Price (SGD$)	Location	Stall Name	Postal Code
Roasted chicken rice	$4.00	Blk 123 Bedok North St 1	Victory chicken rice	460123
Fishball noodles	$3.50	Blk 456 Bedok South Road	All day fishball noodles	460456
Char Siew rice	$3.00	Blk 789 ABC coffeeshop	Best char siew	460789

Data Analysis

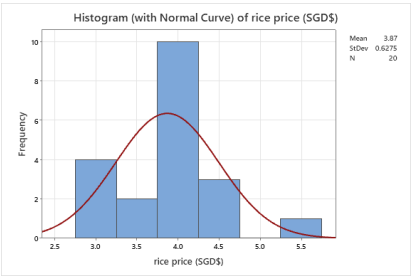
Numerical and graphical summaries

There are $n_1 = 20$ rice dishes and $n_2 = 37$ noodle dishes recorded.

Statistics

Variable	Total Count	Mean	SE Mean	StDev	Variance	Minimum	Q1	Median
rice price (SGD$)	20	3.870	0.140	0.628	0.394	2.800	3.500	4.000

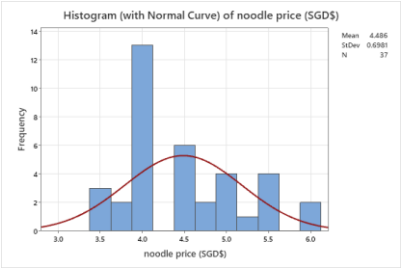
Variable	Q3	Maximum
rice price (SGD$)	4.000	5.500



Statistics

Variable	Total Count	Mean	SE Mean	StDev	Variance	Minimum	Q1	Median
noodle price (SGD$)	37	4.486	0.115	0.698	0.487	3.500	4.000	4.500

Variable	Q3	Maximum
noodle price (SGD$)	5.000	6.000

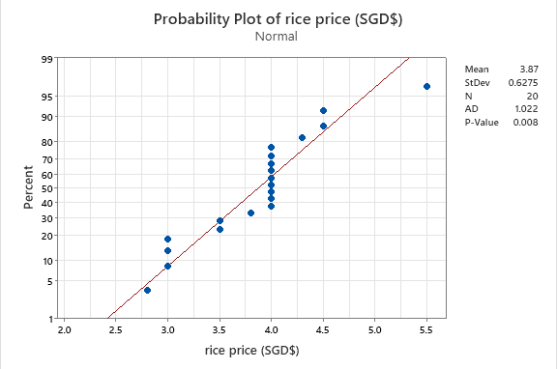


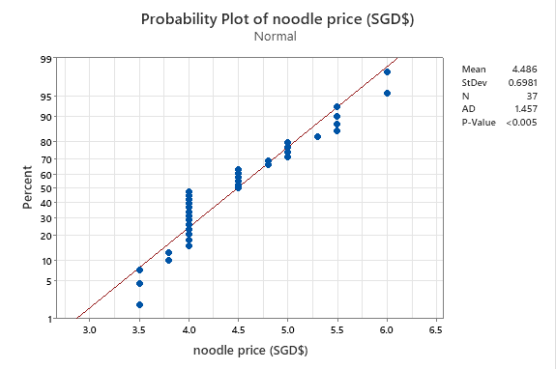
Perform normality test on both samples

Let significance level $\alpha = 5\%$

H_0 : Data follow a normal distribution

H_1 : Data do not follow a normal distribution





Both samples have points not close to line, and P-values for both are $< \alpha = 0.05$, when using Anderson-Darling & Kolmogorov-Smirnov normality tests. Only Ryan-Joiner normality test shows a P-value of > 0.1. Hence, the null hypotheses are rejected for both. Data does not follow a normal distribution for both.

Perform Levene's test to check for equality of variances assuming not normal data

Assume that samples are independent, data is continuous, and homogeneity of variances when using Levene's test.

Let σ_1^2 be the variance of mean price of Chinese rice dishes in Bedok

Let σ_2^2 be the variance of mean price of Chinese noodles dishes in Bedok

$H_0 : \sigma_1^2 / \sigma_2^2 = 1$

$H_1 : \sigma_1^2 / \sigma_2^2 \neq 1$

Method

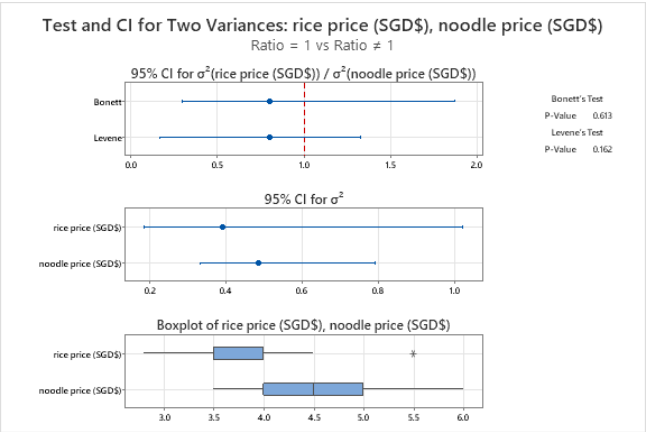
σ_1: standard deviation of rice price (SGD$)
σ_2: standard deviation of noodle price (SGD$)
Ratio: σ_1/σ_2
The Bonett and Levene's methods are valid for any continuous distribution.

Descriptive Statistics

Variable	N	StDev	Variance	95% CI for σ^2
rice price (SGD$)	20	0.628	0.394	(0.186, 1.024)
noodle price (SGD$)	37	0.698	0.487	(0.334, 0.793)

Ratio of Variances

Estimated Ratio	95% CI for Ratio using Bonett	95% CI for Ratio using Levene
0.808084	(0.301, 1.877)	(0.169, 1.332)



Test

Null hypothesis $H_0: \sigma_1^2 / \sigma_2^2 = 1$
Alternative hypothesis $H_1: \sigma_1^2 / \sigma_2^2 \neq 1$
Significance level $\alpha = 0.05$

Test

Method	Statistic	DF1	DF2	P-Value
Bonett	*			0.613
Levene	2.01	1	55	0.162

P-value of Levene's test = 0.162 > α = 0.05. Also, the 95% confidence bound for the ratio of variances is $0.169 < \sigma_1^2 / \sigma_2^2 < 1.332$ which includes 1. Hence, we do not reject null hypothesis. The variances of the mean price of Chinese rice and noodles dishes in Bedok are equal.

Hence, we will select lower tail two sample t-test assuming equal variances.

Lower tail two sample t-test assuming equal variances

Let significance level α = 5%

Let μ_{rice} be the mean price of Chinese rice dishes in Bedok

Let $\mu_{noodles}$ be the mean price of Chinese noodle dishes in Bedok

H_0 : $\mu_{rice} - \mu_{noodles} = 0$; Mean price of Chinese rice dishes is the same as noodle dishes in Bedok

H_1 : $\mu_{rice} - \mu_{noodles} < 0$; Mean price of Chinese rice dishes is cheaper than noodles dishes in Bedok

Method

μ_1: population mean of rice price (SGD$)
μ_2: population mean of noodle price (SGD$)
Difference: $\mu_1 - \mu_2$

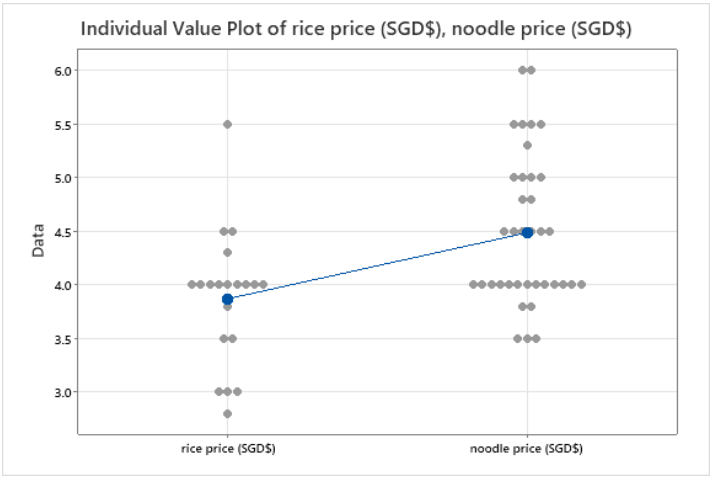
Equal variances are assumed for this analysis.

Descriptive Statistics

Sample	N	Mean	StDev	SE Mean
rice price (SGD$)	20	3.870	0.628	0.14
noodle price (SGD$)	37	4.486	0.698	0.11

Estimation for Difference

Difference	Pooled StDev	95% Upper Bound for Difference
-0.616	0.675	-0.303



Test

Null hypothesis	H_0: $\mu_1 - \mu_2 = 0$
Alternative hypothesis	H_1: $\mu_1 - \mu_2 < 0$

T-Value	DF	P-Value
-3.29	55	0.001

Results Interpretation

Since P-value = 0.001 < α = 0.05, it is rare to obtain a difference of sample mean price of $3.87 - $4.486 = -$0.616 or less, if the population difference is $0. Also, we are 95% confident that the population mean difference falls below -$0.303, which does not include the claimed difference of $0. So, we reject H_0 at α = 5%.

Hence, Chinese rice dishes are cheaper than Chinese noodle dishes in Bedok. Elderly or anyone on a food budget should stick to rice meals instead of noodles.

Nevertheless, there is a possibility of Type I error to conclude that Chinese rice dishes are cheaper than Chinese noodles dishes when in fact it is not.