



JOINT INSTITUTE
交大密西根学院

VE401, Probabilistic Methods in Engineering

Term Project 1:

Analysis of Package Contents

SP2019 Group 14

Group members

Chen Zhiyang 517370910165

Gan Bicheng 517370910049

Hu Yifan 517370910156

Liu Wanzhen 516370910105

Ma Ziqiao 517370910114

Instructed by

Dr. Horst Hohberger

April 8, 2019

UM-SJTU Joint Institute

1 Abstract

This report analyzes and applies the metrology and inspection sampling plan from JTF 1070-2005 to the products in prepackages with fixed content. We explain the sample testing equation and the maximum allowed Type T_1 and T_2 shortfall amounts. Further, we use hypothesis test and OC curve to demonstrate the confidence coefficient of the testing rule. A sample of candies is tested followed relevant procedures and verify the analysis above. Thus, we prove the rationality of the sampling plan and give a deep exploration on its reason.

Key Words hypothesis test,
acceptance sampling,
Metrology and Inspection Sampling Plan,
non-central T distribution,
JTF 1070-2005.

Contents

1	Abstract	1
2	Project Introduction	6
2.1	Problem Description	6
2.2	Project Objectives	6
3	Testing Rules for Mean Net Quantities of a Batch	7
3.1	Objectives	7
3.2	Summary of the Testing Rules [1]	7
3.3	Definition and Notations	8
3.4	Mathematical Interpretation of $\bar{q} \geq (Q_n - \lambda s)$	9
3.4.1	Relations with Confidence Interval	9
3.4.2	Relations with Hypothesis Test	10
3.5	Relations between Confidence Interval and Hypothesis Tests	12
4	Acceptance Sampling for Net Quantities of a Batch	12
4.1	Objectives	12
4.2	Definition and Notation	12
4.3	Calculation of Significance Level	14
4.3.1	Level of Type I Error	14
4.3.2	Level of Type II Error	16
4.4	Derivation of Acceptable Number	19
4.4.1	Significance of α Determining c_{\min}	19
4.4.2	Significance of β Determining c_{\max}	23
4.4.3	Determine c from OC Curves	26
5	Non-central T Distribution	28
5.1	Objectives	28
5.2	Definition and Notations	28
5.3	Derivation of the OC Curve	29
5.4	The OC Curves of the Test in Table 4	33
5.5	Power of the Tests	35
5.6	Probability of Detecting a Shortfall of at Least One Standard Deviation	36
6	Significance Level of Type I Error	37
6.1	Summary of Section 5.3.2 in [1]	37
6.2	Discussions of the Statements	37

7	Tests of a Sample of Prepackaged Food	39
7.1	Objectives	39
7.2	Testing Materials and Apparatus	39
7.2.1	Testing Materials	39
7.2.2	Apparatus	40
7.3	Procedure	40
7.3.1	Measurement of average tare weight	40
7.3.2	Measurement of the actual content deviation	42
7.4	Measurements and Results	42
7.4.1	Measurement of average tare weight	42
7.4.2	Actual content deviation	42
7.5	Analysis	44
8	Conclusion	45
9	References	46
9.1	Works Cited	46
9.2	Figures	46
10	Appendix	46
10.1	Mathematica and Matlab Codes	46

List of Figures

1	A Label Reading [net quantity: 12g]	6
2	A Intuitive Example of Central Limit Theorem [I]	8
3	Lower Bounded Integral Estimation of a Student's T-Distribution	9
4	The Plot of k against σ and Q_n	14
5	The Relationship between α and N	15
6	The Relationship between β and Π , Binomial Approx	17
7	The Relationship between β and Π , Poisson Approx	17
8	The Relationship between β and N , Hypergeometric, at $\Pi = 0.09$.	18
9	The Relationship between c_{\min} and N , Hypergeometric	20
10	The Relationship between c_{\min} and N , Binomial Approx	21
11	The Relationship between c_{\min} and N , Poisson Approx	22
12	The Relationship between c_{\max} and N , Hypergeometric	23
13	The Relationship between c_{\max} and N , Binomial	24
14	The Relationship between c_{\max} and N , Poisson	25
15	OC Curves when $n = 50$	27
16	OC Curves when $n = 80$	27
17	OC Curves when $n = 125$	28
18	Two-sided Integral Estimation of a Non-Central T-Distribution	31
19	Upper Bounded One-sided Integral Estimation of a Non-Central T-Distribution	32
20	Lower Bounded One-sided Integral Estimation of a Non-Central T-Distribution	33
21	The OC Curves of the Test in Table 4	34
22	The OC Curves of the Test in Table 4, $H_0 : \mu \leq Q_n$	34
23	The OC Curves of the Test in Table 4, $H_0 : \mu \geq Q_n$	35
24	The Plot of k against σ and Q_n	38
25	The Relationship between c_{\max} and N , Hypergeometric, $\Pi_1 = 0.12$	39
26	DOUBLEMINT pressed candies	39
27	Electronic Scale	40
28	Measurement of Net Weight	41
29	Measurement of Tare Weight	41
30	The general weight of package	41
31	The boxplot of sample weight	43

List of Tables

1	The Relationship of n and N	7
2	Theoretic and Given Modifying Factors	10
3	α with Corresponding n in Binomial Approx	16
4	α with Corresponding n in Poisson Approx	16
5	The value of β found for $\Pi = 9\%$	18
6	c_{\min} given by Hypergeometric	20
7	c_{\min} given by Binomial	21
8	c_{\min} given by Poisson	22
9	c_{\max} given by Hypergeometric	24
10	c_{\max} given by Binomial	25
11	c_{\max} given by Poisson	26
12	Main Labels on the Package	40
13	Main Parameters of the Scale	40
14	Net and Tare Weight of the Samples	42
15	All Measurements of Tare Weights	42
16	Calculation of Actual Content Deviation	43

2 Project Introduction

2.1 Problem Description

According to JJF 1070-2005, products that are packaged with packing materials or containers before sale, and has predetermined measures or numerical counts are prepackaged products [1]. Prepackaged food with label is commonly seen in daily life, including wrapped candy bars, boxed cookies and bottled drinks. When coming across a box of DOUBLEMINT candy whose label reads [net quantity: 12g], you may have the experience of weighing the item in the hand with doubt.



Figure 1: A Label Reading [net quantity: 12g]

Thankfully, most governments have developed high awareness towards labelling issues and put forward series of detailed regulations. For example, the FD&C Act provides specifically that any food, drug, or cosmetic "shall be deemed misbranded" if its label fails to contain "an accurate statement of the quantity of contents in terms of weight, measure, or numerical count" [2]. In general, any prepackaged food must be labelled with conspicuousness the nominal content that the product contains.

Statistically, it then becomes a problem how people decide whether to accept or not a large batch of prepackaged product, where a batch usually contains incredibly many unit products that it's impossible to check them one-by-one. On top of that, Chinese government develop a system of rules of metrological testing for net quantities of products, which the project is interested in.

2.2 Project Objectives

The project is based on *Rules of metrological testing for net quantities of products in prepackages with fixed content*, issued by Chinese government in 2005 [1].

Specifically, we are interested in the methods and rules that Chinese government predetermined to:

- measure the net quantity of a batch of products;
- control the type I error (rejecting an acceptable lot) and type II error (failing to reject an unacceptable lot).

We will then conduct series of tests with accord to the document, in order to verify whether or not the actual quantity of a product sample corresponds to its labelled contents.

3 Testing Rules for Mean Net Quantities of a Batch

3.1 Objectives

This section is oriented to question i) and the pre-question summary.

The objective of this section is to:

- summarize the testing rules put forward by the document.
- verify the validity of the modifying factors λ that the document preset.

After the section, readers should be able to comprehend:

- the basic rules of net quantity measurement
- the relationship of the modifying factor with confidence interval and hypothesis tests

In the process, readers will be shown a series of figures to visualize the concept.

3.2 Summary of the Testing Rules [1]

The general testing goal is to ensure that the average actual quantity of a batch of products in prepackages with fixed content should be greater or equal to its nominal quantity.

1. When the batch number $1 \leq N \leq 10$:

We apply general investigating (all unit products are to be tested). The batch is accepted only if all the products have a nominal quantity greater or equal to nominal quantity.

2. When the batch number $N \geq 11$: We apply sampling (only some unit products in the batch are to be tested). The relationship of sample size and batch number is given by Table 1:

Batch Number (N)	Sample Size (n)	Acceptance Number (c)
11 ~ 50	10	0
51 ~ 99	13	1
100 ~ 500	50	3
501 ~ 3200	80	5
3200 ~ ∞	125	7

Table 1: The Relationship of n and N

The sample tested should satisfy all the following statements to get accepted:

- The sample mean \bar{q} should be greater or equal to the difference between the nominal quantity and the product of sample standard deviation and modifying factor $\lambda = t_{0.995}/\sqrt{n}$.

$$\bar{q} \geq Q_n - \lambda s$$

- The number of type T_1 shortfalls should be less than or equal to the preset acceptance number (Table 1).
- No type T_2 shortfall is detected.

3.3 Definition and Notations

Central Limit Theorem [3] The central limit theorem implies that random errors in physical quantity measurements tend to be normal because each individual contributing random errors add up to be normal.

Based on this theorem, in this project, we approximate each individual net quantity by a normal distribution with mean μ and variance σ^2 .

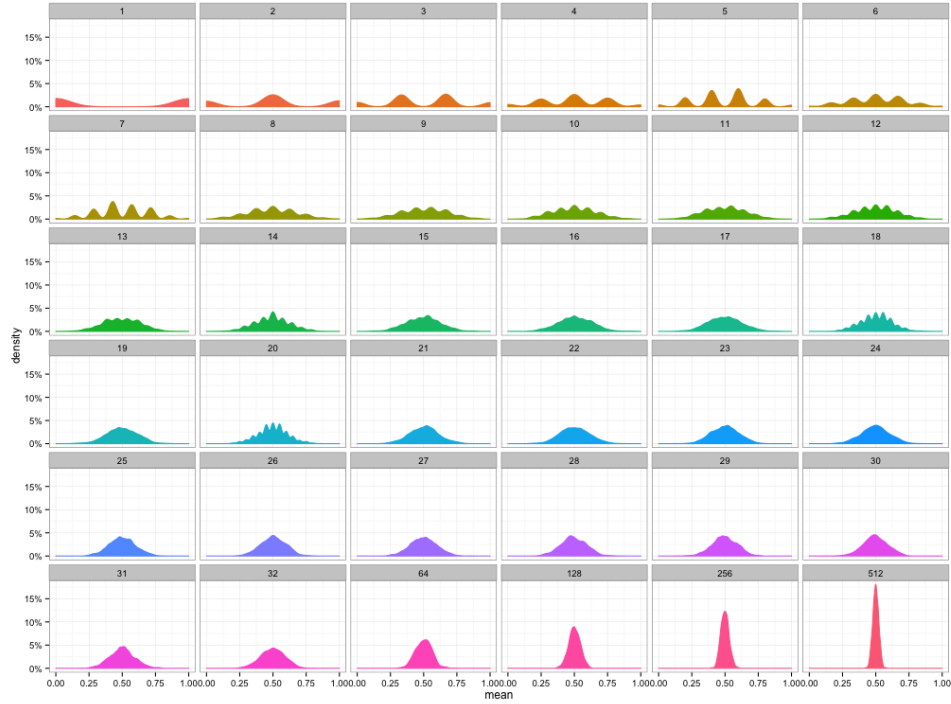


Figure 2: A Intuitive Example of Central Limit Theorem [1]

Confidence Interval [3] Based on some sampling results, we may conclude with some level of certainty that a population parameter lies in some interval, and the interval is known as confidence interval.

In this section, we are particular interested in **lower bounded one-sided confidence interval** $[L, \infty)$, where with a $100(1 - \alpha)\%$ lower confidence bound L for population parameter θ ,

$$P[L \leq \theta] = 1 - \alpha$$

3.4 Mathematical Interpretation of $\bar{q} \geq (Q_n - \lambda s)$

We need to satisfy the rule that the actual average quality of batch quantitative packaging of goods μ is greater or equal to the net quantity Q_n .

3.4.1 Relations with Confidence Interval

We are going to explain its relation to confidence interval from the interval estimation of mean \bar{Q} .

In terms of a sample following normal distribution, the value of σ is unclear, thus needs to be estimated by S , the sample standard deviation. To find an one-sided 99.5% confidence interval of the mean, we need the Student's T-Distribution:

$$T_{n-1} = \frac{\bar{Q} - \mu}{S/\sqrt{n}}$$

We use the similar notation defined in the slides [3]. Let $0 < \alpha \leq 1$ we define $t_{\alpha,n} \geq 0$ by

$$\int_{t_{\alpha,n}}^{\infty} = \alpha/2$$

This can be visualized by the following figure.

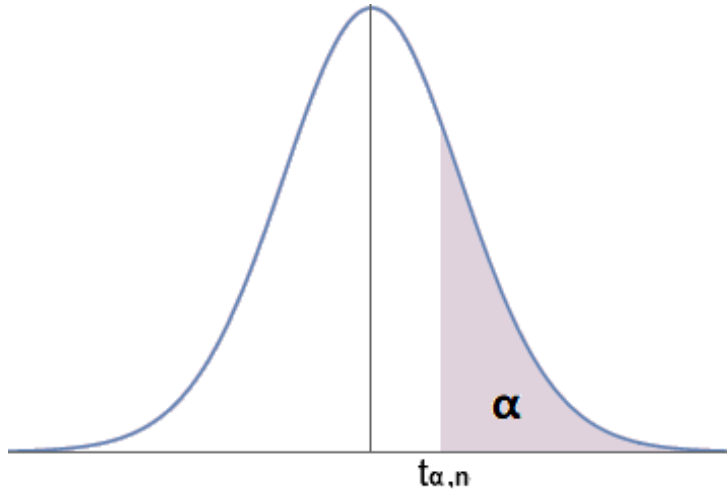


Figure 3: Lower Bounded Integral Estimation of a Student's T-Distribution

We now derive the one-sided confidence interval of the sample mean.

Theorem A $100(1 - \alpha)\%$ upper confidence bound on μ is given by $\bar{Q} + \frac{st_{\alpha,n-1}}{\sqrt{n}}$.

The proof is simply as follows. From Student T-Distribution,

$$P\left[\frac{\bar{Q} - \mu}{S/\sqrt{n}} \geq -t_{\alpha, n-1}\right] = 1 - \alpha$$

$$P\left[\mu \leq \bar{Q} + \frac{s \cdot t_{\alpha, n-1}}{\sqrt{n}}\right] = 1 - \alpha$$

We noticed a term t_α defined in the Table 4 in [1]. Through above discussion, we find out the meaning of this notation.

Notation

$$t_{1-\alpha} = \frac{t_{\alpha, n-1}}{\sqrt{n}} \quad (1)$$

Using notation (1), we now have:

$$P[\mu \leq \bar{Q} + st_\alpha] = 1 - \alpha$$

which indicates that for a batch of products with $\mu = Q_n$,

$$P[Q_n \leq \bar{Q} + st_\alpha] = 1 - \alpha$$

$$P[\bar{Q} \geq Q_n - st_\alpha] = 1 - \alpha$$

We have 99.5% confidence that the batch will be accepted, if the sample mean \bar{q} satisfies that:

$$\bar{q} \geq Q_n - st_\alpha$$

For verification that our assumption is correct, we calculated the theoretic modifying factor λ for given n and compare with given n .

Sample Size n	Theoretic λ	Given λ in Table 4
10	1.027688233	1.028
13	0.847176855	0.848
50	0.379002443	0.379
80	0.295105589	0.295
125	0.233987509	0.234

Table 2: Theoretic and Given Modifying Factors

We observed that the values correspond.

3.4.2 Relations with Hypothesis Test

Relations with Neyman-Pearson Hypothesis Test In a Neyman-Pearson Null Hypothesis Test, we basically want to make a decision between two discrete hypothesis by accepting one of them.

To explain this table in the view of Neyman-Pearson Hypothesis Test, we set out the hypotheses as:

$$H_0 : \mu = Q_n$$

$$H_1 : |\mu - Q_n| \geq \delta$$

and try to choose between them.

To set up the test, we need to select a test statistic and define a critical region. From Section 3.4.1, with a probability of $1 - \alpha$,

$$T_{n-1} \geq -t_{\alpha, n-1}$$

If H_0 is true,

$$P\left[\frac{\bar{Q} - Q_n}{S/\sqrt{n}} < -t_{\alpha, n-1}\right] = \alpha$$

giving the critical region

$$\bar{q} < Q_n - \frac{st_{-\alpha, n-1}}{\sqrt{n}}$$

Therefore, the sample data with $\bar{q} \geq Q_n - \lambda = Q_n - \frac{st_{-\alpha, n-1}}{\sqrt{n}}$ falls out of the critical region, and we hence fail to reject H_0 .

Since $\alpha = 0.005$, the chance of this being an error is only 0.005. Note that this implies that a lot is acceptable unless proven otherwise.

Relations with Fisher's Hypothesis Test In a Fisher's Null Hypothesis Test, we basically look for evidence against some hypothesis H_0 . From [3], we know that the P Value (level of significance) is the upper bound of the probability obtaining the data when the hypothesis is true, i.e.

$$P = P[D \mid H_0]$$

To explain this table in the view of Fisher's Null Hypothesis Test, we set out the null hypothesis as:

$$H_0 : \mu = Q_n$$

and try to find evidence against it.

In the following we will verify that the inequality $\bar{q} \geq (Q_n - \lambda s)$ ensures that a batch with $\mu = Q_n$ will be rejected at no more than 0.005 level of significance, as is required in [1, section 5.3.2.1].

We assume the same definition of t_α from Section 3.4.1. From definition, the P value can be calculated by the following one-sided interval:

$$\begin{aligned} P[D \mid H_0] &= P[\bar{Q} \geq Q_n - t_\alpha s \mid \mu = Q_n] \\ &= P[\bar{Q} \geq \mu - t_\alpha s] \\ &= P[\bar{Q} \geq \mu - \frac{t_{\alpha, n-1} s}{n}] \\ &= P\left[\frac{\bar{Q} - \mu}{S/\sqrt{n}} \geq -t_{\alpha, n-1}\right] \\ &= 1 - \alpha \end{aligned}$$

In [3, Table 4], it's preset that $\alpha = 0.995$. Hence the P value is given by $P = 0.005$, which is small enough to reject H_0 .

Therefore, there is evidence that μ is greater than Q_n if we received a sample data with $\bar{q} \geq Q_n - \lambda$.

3.5 Relations between Confidence Interval and Hypothesis Tests

From the above discussion of the two-tailed null hypothesis

$$H_0 : \mu = Q_n$$

We conclude that:

- Neyman-Pearson:

\bar{q} lies in the critical region of $\alpha \Leftrightarrow$ null value Q_n does not lie in the $100(1 - \alpha)\%$ one-sided upper bounded confidence interval for μ .

- Fisher's:

H_0 will be rejected at α level of significance \Leftrightarrow null value Q_n does not lie in the $100(1 - \alpha)\%$ one-sided upper bounded confidence interval for μ .

4 Acceptance Sampling for Net Quantities of a Batch

4.1 Objectives

This section is oriented to question ii).

The objective of this section is to:

- calculate the significance of producer's and customer's risks;
- verify the validity of the acceptance numbers c that the document preset.

After the section, readers should be able to comprehend:

- how risks are controlled through the significance level;
- how the acceptance numbers were derived under the restrictions of producer's risk α , customer's risk β and type T_2 shortfall.

In the process, readers will be shown a series of figures to visualize the concept.

4.2 Definition and Notation

Acceptance Sampling [3] In acceptance sampling, we apply the following notations:

We denote the batch number by N and sample number by n still. For the actual but unknown proportion of products with T_1 shortfall, we use the notation Π , and ideally we accept the batch if this proportion is less than Π_0 . To lower the costumer's risk, we set another "barely acceptable" proportion Π_1 .

Upon that, we make two one-sided hypotheses:

$$H_0 : \Pi \leq \Pi_0$$

$$H_1 : \Pi > \Pi_1$$

In sample testing, we define acceptance number c that if the number of defectives is less than or equal to c , we accept the batch of products.

Ignorance of T_2 Shortfall In the fifth and sixth column, the testing method is similar to the process of acceptance sampling, except that "defective items" are divided into T_1 shortfall and T_2 shortfall.

In the following section, we will approximate the test by the acceptance sampling process, by assuming that no T_2 shortfall happens. To validate this assumption, we will prove that T_2 shortfalls are ignorable when the Q_n is large and σ is small.

From Central Limit Theorem, we assumed that the net quantity Q of a unit product follows a normal distribution with variance μ and standard deviation σ . We further assume that $\mu \approx Q_n$.

The probability of detecting an T_1 shortfall is given by

$$\begin{aligned} P[Q_n - 2T \leq Q < Q_n - T] &= P\left[\frac{Q_n - 2T - \mu}{\sigma} \leq \frac{Q - \mu}{\sigma} < \frac{Q_n - T - \mu}{\sigma}\right] \\ &= P\left[\frac{Q_n - 2T - \mu}{\sigma} \leq Z < \frac{Q_n - T - \mu}{\sigma}\right] \\ &= \Phi\left(\frac{Q_n - T - \mu}{\sigma}\right) - \Phi\left(\frac{Q_n - 2T - \mu}{\sigma}\right) \\ &= \Phi\left(\frac{-T}{\sigma}\right) - \Phi\left(\frac{-2T}{\sigma}\right) \end{aligned}$$

The probability of detecting an T_2 shortfall is given by

$$\begin{aligned} P[Q < Q_n - 2T] &= P\left[\frac{Q - \mu}{\sigma} < \frac{Q_n - 2T - \mu}{\sigma}\right] \\ &= P\left[Z < \frac{Q_n - 2T - \mu}{\sigma}\right] \\ &= \Phi\left(\frac{-2T}{\sigma}\right) \end{aligned}$$

We denote the ratio $k := \frac{P[Q < Q_n - 2T]}{P[Q_n - 2T \leq Q < Q_n - T]}$, which measures the probability of T_2 shortfall compare to T_1 .

Using **Mathematica**, we plot k against σ (ranging between 0 and 1) and Q_n (ranging between 0 and 10). The plot is given by:

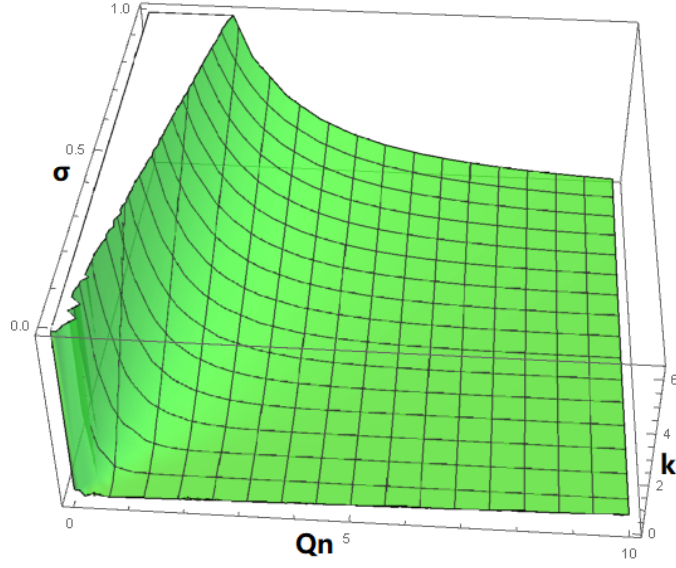


Figure 4: The Plot of k against σ and Q_n

Reading from Figure 4 we directly note that T_2 shortfalls are ignorable when the Q_n is large and σ is small. For example, we fix $\sigma = 0.4$ and $\mu = Q_n = 8$, we can calculate the k value is only:

$$k = 0.00444795$$

However, situations when the batch has a large σ and small Q_n will be discussed at the end of this section.

4.3 Calculation of Significance Level

4.3.1 Level of Type I Error

The significance Level of Type I Error is known as producer's risk α , which highlights the probability of rejecting an acceptable batch by mistake. From definition,

$$\begin{aligned} \alpha &= P[\text{reject } H_0 \mid H_0 \text{ is true}] \\ &= P[d \geq c \mid \Pi \leq \Pi_0] \\ &\leq P[d \geq c \mid \Pi = \Pi_0] \\ &= \sum_{d \geq c} \frac{\binom{r}{d} \binom{N-r}{n-d}}{\binom{N}{n}} \end{aligned}$$

In [1, Section 5.3.2.2], the document preset that in the acceptance sampling,

$$\Pi_0 = 0.0025$$

which we can apply in the above formula and find the value of α .

Hypergeometric Distribution From above discussion, we have derived a formula:

$$\alpha = \sum_{d \geq c} \frac{\binom{r}{d} \binom{N-r}{n-d}}{\binom{N}{n}}$$

where c is a function of n , n is a function of N , and

$$r = N\Pi = N\Pi_0$$

Hence, we derive a function of α against N :

$$\alpha(N) = \sum_{d \geq c} \frac{\binom{N\Pi_0}{d} \binom{N(1-\Pi_0)}{n(N)-d}}{\binom{N}{n(N)}} \quad (2)$$

Using **Matlab**, we can plot the function as follows.

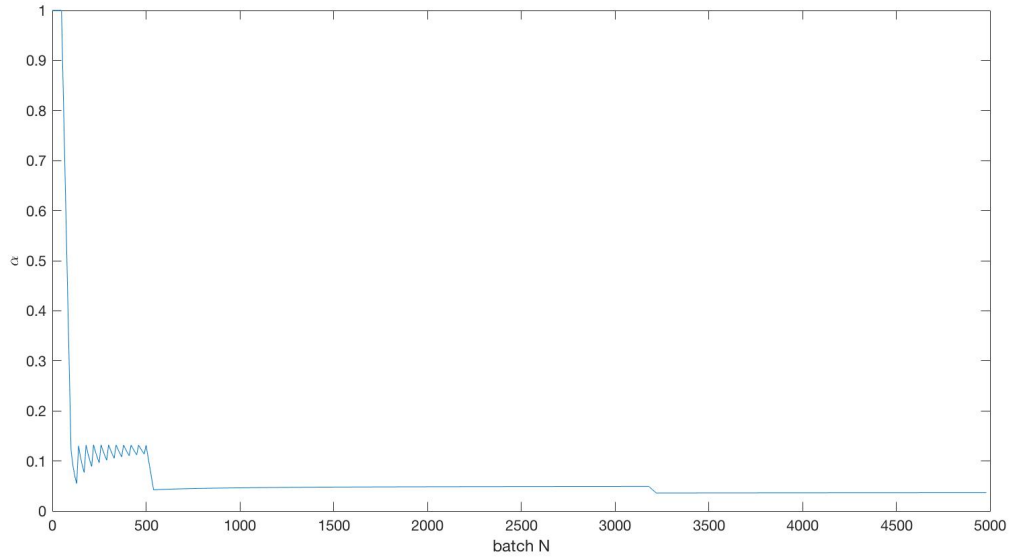


Figure 5: The Relationship between α and N

Binomial Distribution We can use a Binomial Distribution as approximation to eliminate large N in

$$\alpha = \sum_{d \geq c} \frac{\binom{r}{d} \binom{N-r}{n-d}}{\binom{N}{n}}$$

By $p = \frac{r}{N} = \Pi = \Pi_0$, we have

$$\alpha \approx \sum_{d \geq c} \binom{n}{d} \Pi_0^d (1 - \Pi_0)^{n-d}$$

Plugging in the given data, we found

n	10	13	50	80	125
c	0	1	3	5	7
α (binomial)	1	0.28045161	0.12937796	0.05036913	0.03815242

Table 3: α with Corresponding n in Binomial Approx

Poisson Distribution We can use a Poisson Distribution as approximation to eliminate large N in

$$\alpha = \sum_{d \geq c} \frac{\binom{r}{d} \binom{N-r}{n-d}}{\binom{N}{n}}$$

By $k = \frac{nr}{N} = n\Pi = n\Pi_0$, we have

$$\alpha \approx \sum_{d \geq c} \frac{(n\Pi_0)^d e^{-\Pi_0}}{d!}$$

Plugging in the given data, we found

n	10	13	50	80	125
c	0	1	3	5	7
α (poisson)	1	0.27747265	0.13153233	0.05265302	0.04020872

Table 4: α with Corresponding n in Poisson Approx

We found the value of α unacceptably large for small n . This will be discussed in Section 6, question iv).

4.3.2 Level of Type II Error

The significance Level of Type II Error is known as customer's risk α , which highlights the probability of accepting an unacceptable batch by mistake. From definition,

$$\begin{aligned} \beta &= P[\text{failing to reject } H_0 \mid H_1 \text{ is true}] \\ &= P[d \leq c \mid \Pi > \Pi_1] \\ &= \sum_{d \leq c} \frac{\binom{r}{d} \binom{N-r}{n-d}}{\binom{N}{n}} \\ &= \sum_{d \leq c} \frac{\binom{n\Pi}{d} \binom{(1-\Pi)N}{n-d}}{\binom{N}{n}} \end{aligned}$$

We know that $\beta(\Pi)$ is a function of Π , and it is meaningless to analyse N in β when Π is unknown, therefore, we apply approximations.

Binomial Distribution We can use a Binomial Distribution as approximation to eliminate large N by $p = \frac{r}{N} = \Pi = \Pi_0$:

$$\beta \approx \sum_{d \leq c} \binom{n}{d} \Pi^d (1 - \Pi)^{n-d}$$

Using `Matlab`, we can plot the function β v.s. Π :

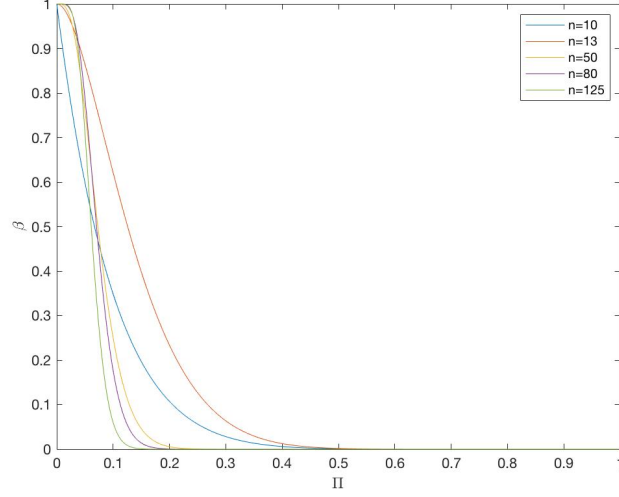


Figure 6: The Relationship between β and Π , Binomial Approx

Poisson Distribution We can use a Poisson Distribution as approximation to eliminate large N by $k = \frac{nr}{N} = n\Pi = n\Pi_0$:

$$\beta \approx \sum_{d \leq c} \frac{(n\Pi)^d e^{-\Pi}}{d!}$$

Using `Matlab`, we can plot the function β v.s. Π :

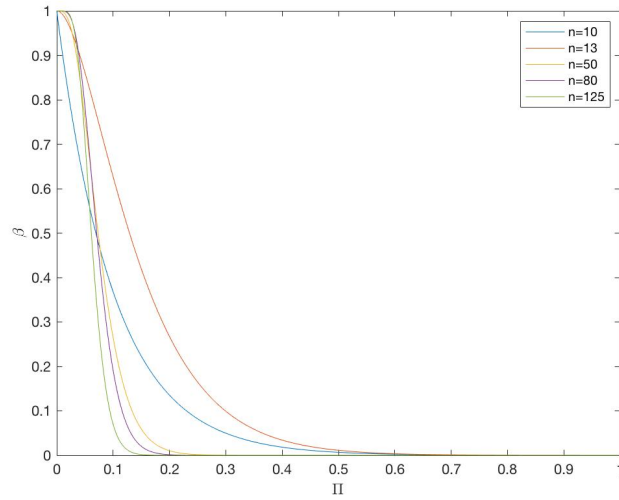


Figure 7: The Relationship between β and Π , Poisson Approx

In [1, Section 5.3.3], the document preset that in the acceptance sampling, we want to verify the value of β when

$$\Pi = 0.09$$

Given the value of Π , we can now calculate the value of β for each n using Poisson and Binomial Distribution Approximation, and plot the function of β against N for the original Hypergeometric formula, by plugging $\Pi = 0.09$ in the previously derived formulae:

Hypergeometric:

$$\beta = \sum_{d \leq c} \frac{\binom{n\Pi}{d} \binom{(1-\Pi)N}{n-d}}{\binom{N}{n}}$$

Binomial:

$$\beta \approx \sum_{d \leq c} \binom{n}{d} \Pi^d (1 - \Pi)^{n-d}$$

Poisson:

$$\beta \approx \sum_{d \leq c} \frac{(n\Pi)^d e^{-\Pi}}{d!}$$

The results are as follows.

n	10	13	50	80	125
c	0	1	3	5	7
β (poisson)	0.78	0.96	0.96	0.98	0.99
β (binomial)	0.78	0.96	0.96	0.98	0.99

Table 5: The value of β found for $\Pi = 9\%$.

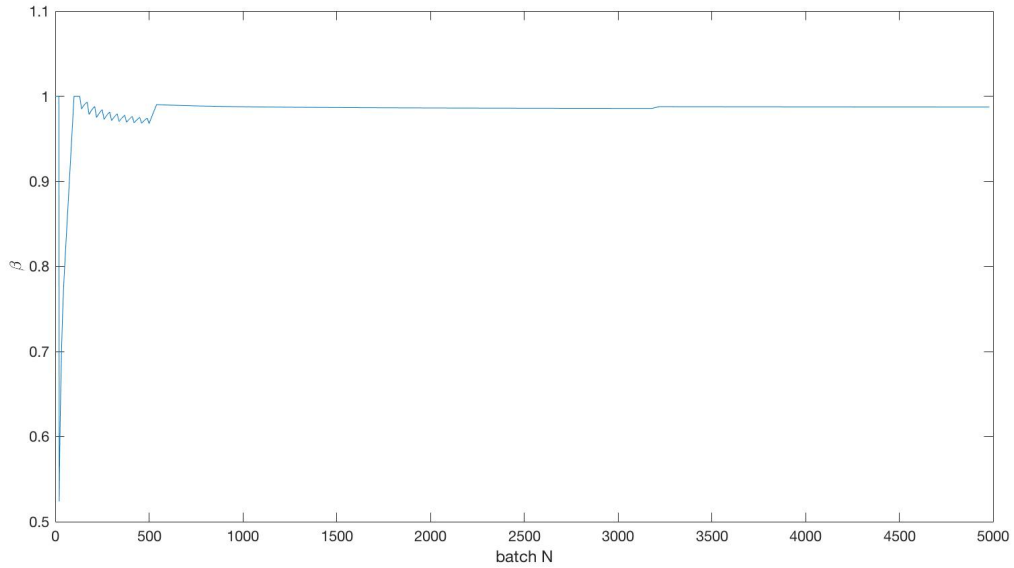


Figure 8: The Relationship between β and N , Hypergeometric, at $\Pi = 0.09$.

4.4 Derivation of Acceptable Number

In this subsection, we first try to find the minimum and maximum of the acceptance number c , and shoot some deviations we found. Then, we apply a new method using OC Curves to find the adequate c , and give illustrative examples.

4.4.1 Significance of α Determining c_{\min}

In [1, Section 5.3.2.2], the document preset that in the acceptance sampling,

$$\Pi_0 = 0.0025$$

$$\alpha \leq 0.05$$

Plugging these parameters into the equalities that we derived in Section 4.3.1, we can find out the minimum integer number c as the acceptance number.

In the following paragraphs, we basically want to plot a function of c_{\min} versus N , and read out the c from the plot.

Hypergeometric Distribution The original value of α is found from Hypergeometric distribution:

$$\alpha = \sum_{d \geq c} \frac{\binom{r}{d} \binom{N-r}{n-d}}{\binom{N}{n}} \leq 0.05$$

We want to find the function of c_{\min} v.s. N , but the function is complicated and in fact it takes **Matlab** a few minutes for calculations. Mathematically, we limited the parameters:

1. The value of d .

To make sure that the codes run well, we set restrictions that:

$$\max(0, n - N + r) \leq d \leq \min(n, r)$$

2. The value of r .

The number r should be integer. For accuracy, we set:

$$r = \text{round}(N\Pi_0)$$

3. The value of c should be integer.

The plot is given as:

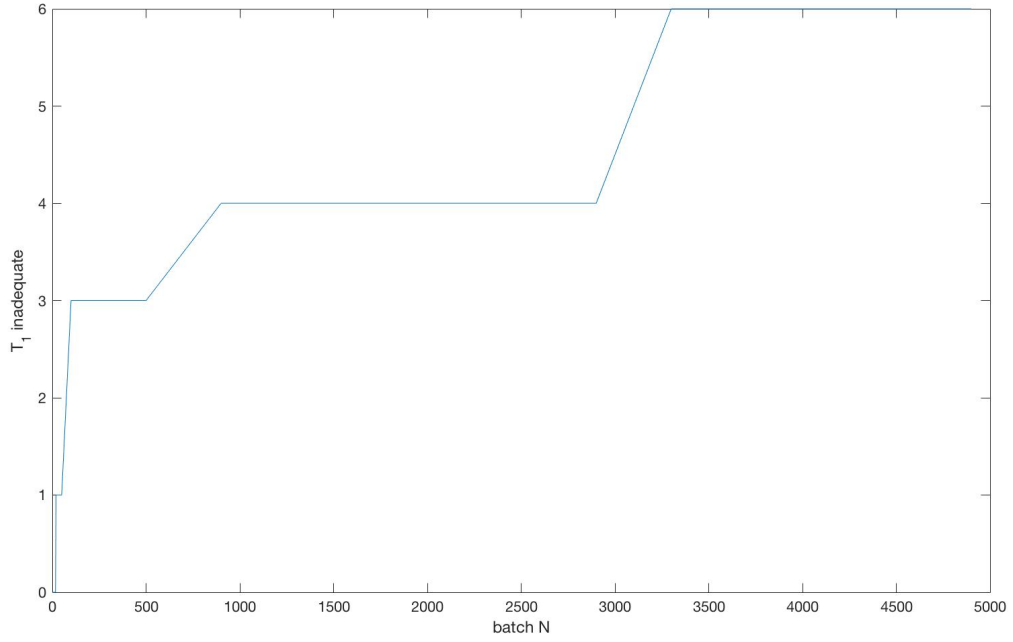


Figure 9: The Relationship between c_{\min} and N , Hypergeometric

From the plot, we read out that:

Sample Size n	Given Acceptance Number c	Theoretic c_{\min} from Hypergeometric
10	0	0
13	1	1
50	3	3
80	5	4
125	7	6

Table 6: c_{\min} given by Hypergeometric

We observed a good correspondence for smaller N , as all the c_{\min} are smaller than the given c .

Binomial Distribution We also have the Binomial approximated α :

$$\alpha \approx \sum_{d \geq c} \binom{n}{d} \Pi_0^d (1 - \Pi_0)^{n-d}$$

which gives good approximations for larger N .

Doing the same thing, we obtained the following Figure from **Matlab**.

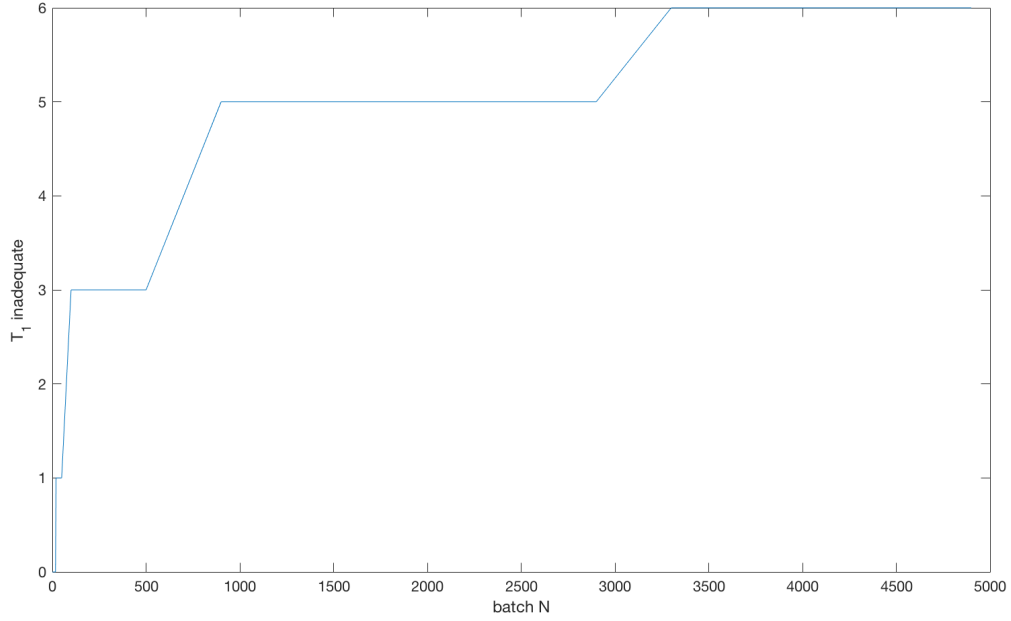


Figure 10: The Relationship between c_{\min} and N , Binomial Approx

From the plot, we read out that:

Sample Size n	Given Acceptance Number c	Theoretic c_{\min} from Binomial
10	0	0
13	1	1
50	3	3
80	5	5
125	7	6

Table 7: c_{\min} given by Binomial

We observed a better correspondence for smaller N , as all the c_{\min} are smaller than the given c , and for $n = 80$, the calculated c coincide with given ones.

Poisson Distribution We also have the Poisson approximated α :

$$\alpha \approx \sum_{d \geq c} \frac{(n\Pi_0)^d e^{-\Pi_0}}{d!}$$

which gives good approximations for larger N .

Doing the same thing, we obtained the following Figure from **Matlab**.

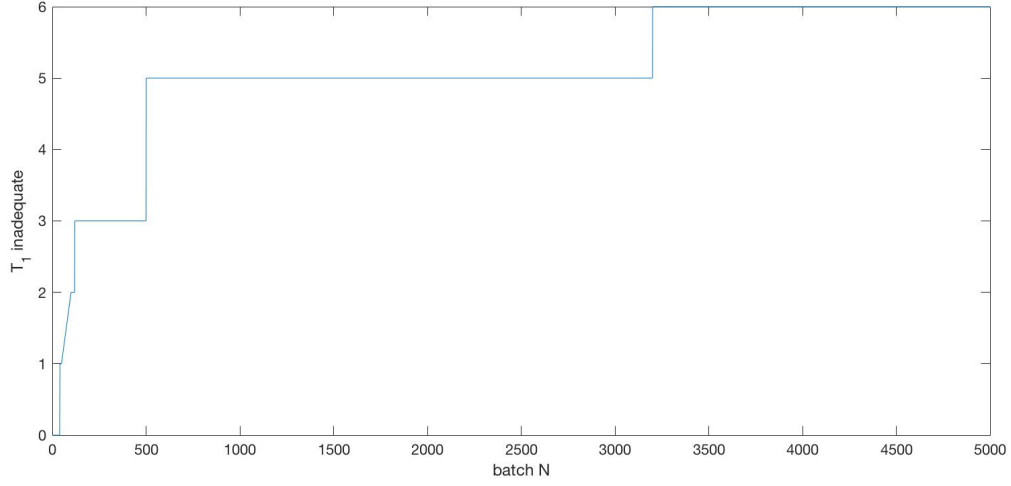


Figure 11: The Relationship between c_{\min} and N , Poisson Approx

From the plot, we read out that:

Sample Size n	Given Acceptance Number c	Theoretic c_{\min} from Poisson
10	0	0
13	1	1
50	3	2,3
80	5	5
125	7	6

Table 8: c_{\min} given by Poisson

We observed a good correspondence for smaller N , as all the c_{\min} are smaller than the given c , but not as good as Binomial.

4.4.2 Significance of β Determining c_{\max}

In [1, Section 5.3.3], the document preset that in the acceptance sampling,

$$\Pi = 0.09$$

$$1 - \beta \geq 0.9 \Rightarrow \beta \leq 0.1$$

Plugging these parameters into the equalities that we derived in Section 4.3.1, we can find out the maximum integer number c as the acceptance number.

In the following paragraphs, we basically want to plot a function of c_{\max} versus N , and read out the c from the plot.

Hypergeometric Distribution The original value of α is found from Hypergeometric distribution:

$$\beta = \sum_{d \leq c} \frac{\binom{n\Pi}{d} \binom{(1-\Pi)N}{n-d}}{\binom{N}{n}} \leq 0.1$$

We want to find the function of c_{\max} v.s. N , similarly, the plot is given as:

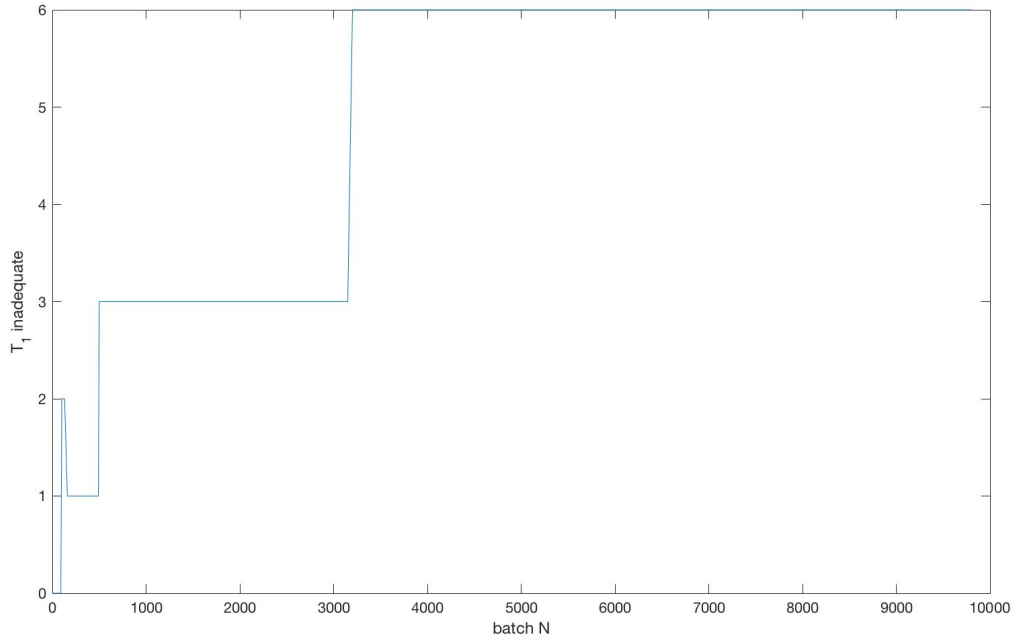


Figure 12: The Relationship between c_{\max} and N , Hypergeometric

From the plot, we read out that:

Sample Size n	Given Acceptance Number c	Theoretic c_{\min} from Hypergeometric
10	0	0
13	1	1,2
50	3	1
80	5	3
125	7	6

Table 9: c_{\max} given by Hypergeometric

The correspondence for larger N is bad, because the maximum c calculated turned out to be less than the given c .

Binomial Distribution Similarly, we repeat the process with Binomial approximation:

$$\beta \approx \sum_{d \leq c} \binom{n}{d} \Pi^d (1 - \Pi)^{n-d}$$

The plot is given by

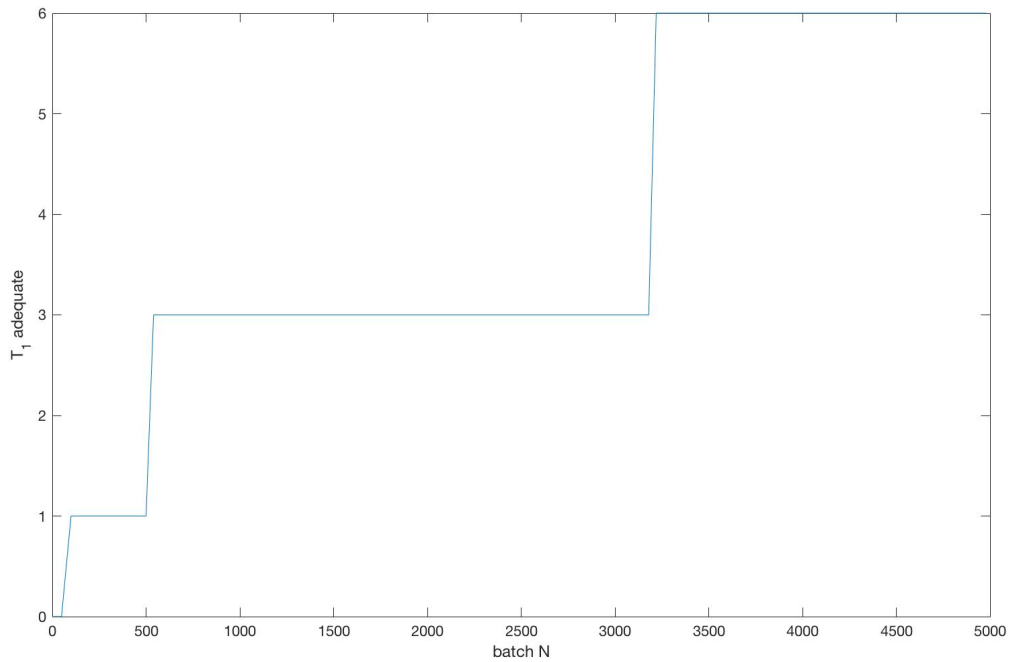


Figure 13: The Relationship between c_{\max} and N , Binomial

From the plot, we read out that:

Sample Size n	Given Acceptance Number c	Theoretic c_{\min} from Binomial
10	0	0
13	1	1
50	3	1
80	5	3
125	7	6

Table 10: c_{\max} given by Binomial

The correspondence for larger N is still bad, because the maximum c calculated turned out to be less than the given c .

Poisson Distribution

$$\beta \approx \sum_{d \leq c} \frac{(n\Pi)^d e^{-\Pi}}{d!}$$

The plot is given by

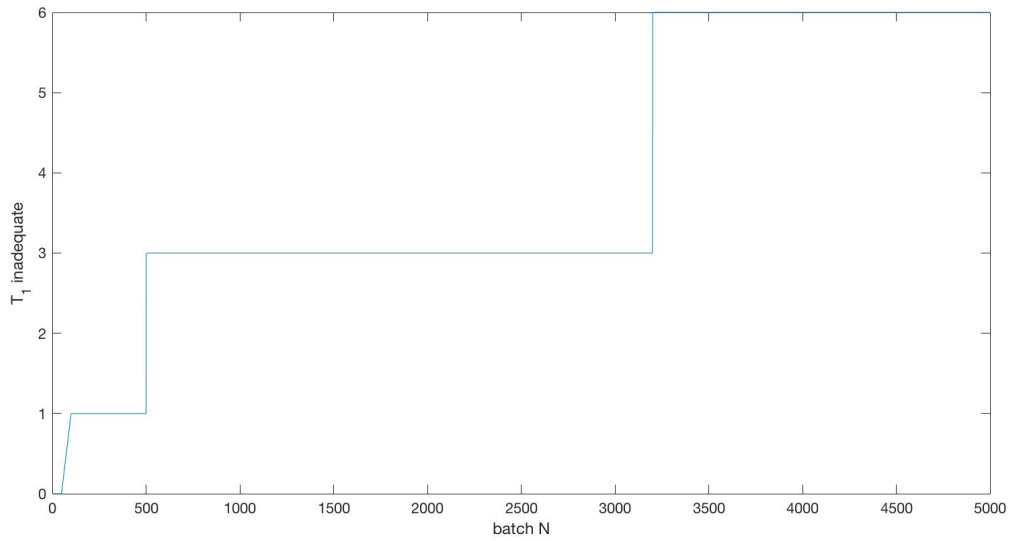


Figure 14: The Relationship between c_{\max} and N , Poisson

From the plot, we read out that:

Sample Size n	Given Acceptance Number c	Theoretic c_{\min} from Poisson
10	0	0
13	1	1
50	3	1
80	5	3
125	7	6

Table 11: c_{\max} given by Poisson

The correspondence for larger N is still bad, because the maximum c calculated turned out to be less than the given c .

We found that α gives almost perfect lower bound of c , but β fails to correspond to c for the larger N . To look into this problem, we apply OC Curves for analysis.

4.4.3 Determine c from OC Curves

In the previous examples, we basically fix the value of Π and look for the relationship of bounds of c with N . We found that for $n \geq 50$, the results are not corresponding. Now, we would like to fix the value of n , and plot the OC Curves with abscissa Π by approximations.

The formula for the OC Curves is given by

$$\begin{aligned}
 P[\text{accpet } H_0] &= P[d \leq c] \\
 &= \sum_{d \leq c} \frac{\binom{N\Pi}{d} \binom{N(1-\Pi)}{n-d}}{\binom{N}{n}} \\
 &\approx \sum_{d \leq c} \frac{(n\Pi)^d e^{-\Pi}}{d!} \quad (\text{Poisson Approx})
 \end{aligned}$$

We apply Poisson approximation since $n \geq 50$ is large enough, and we can plot the OC Curves with testing parameter c . Since α and β require that

$$\Pi_0 = 0.025, \quad \alpha \leq 0.05$$

$$\Pi = 0.09, \quad \beta \leq 0.1$$

On the OC Curves, we make auxiliary lines for α :

$$P = 1 - \alpha = 0.995$$

$$\Pi = \Pi_0 = 0.025$$

and for β :

$$P = \beta = 0.1$$

$$\Pi = 0.09$$

The OC Curves for each given n are plotted as follows by **Matlab**.

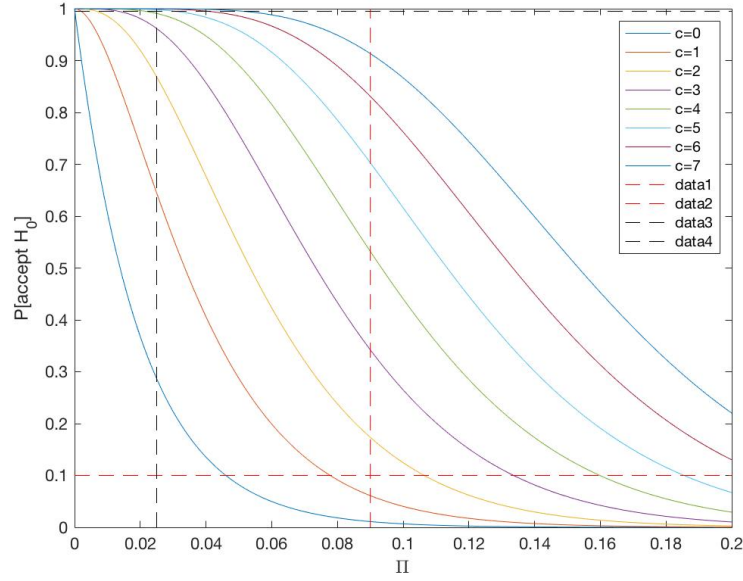


Figure 15: OC Curves when $n = 50$

For $n = 50$, we read from the plot that α requires $c > 4$ and β requires $c < 2$. Clearly it's not possible so we seek for middle solution. We found that the customer's risk is unreasonably high for larger c , and the producer's risk gets larger for smaller c . Therefore, we choose a middle solution $c = 3$.

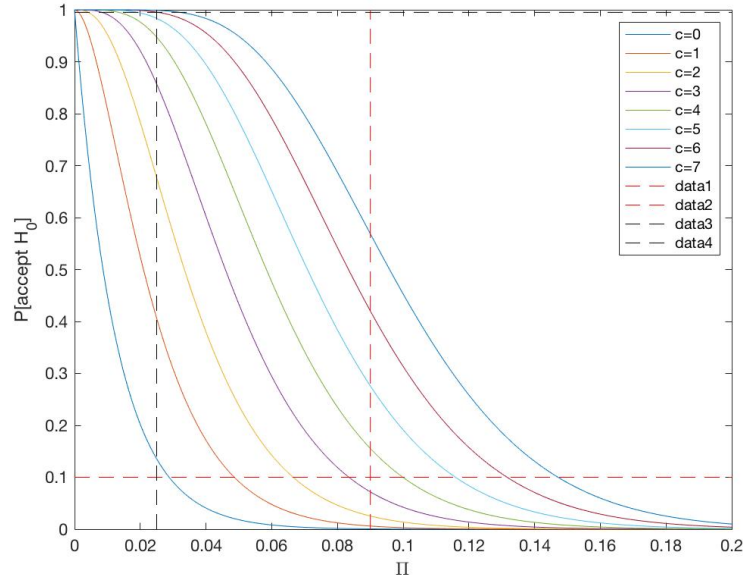


Figure 16: OC Curves when $n = 80$

For $n = 80$, we read from the plot that α requires $c > 5$ and β requires $c < 4$. Clearly it's not possible so we seek for middle solution. We found that the customer's risk is unreasonably high for larger c , and the producer's risk gets larger for smaller c . Therefore, we choose a middle solution $c = 5$.

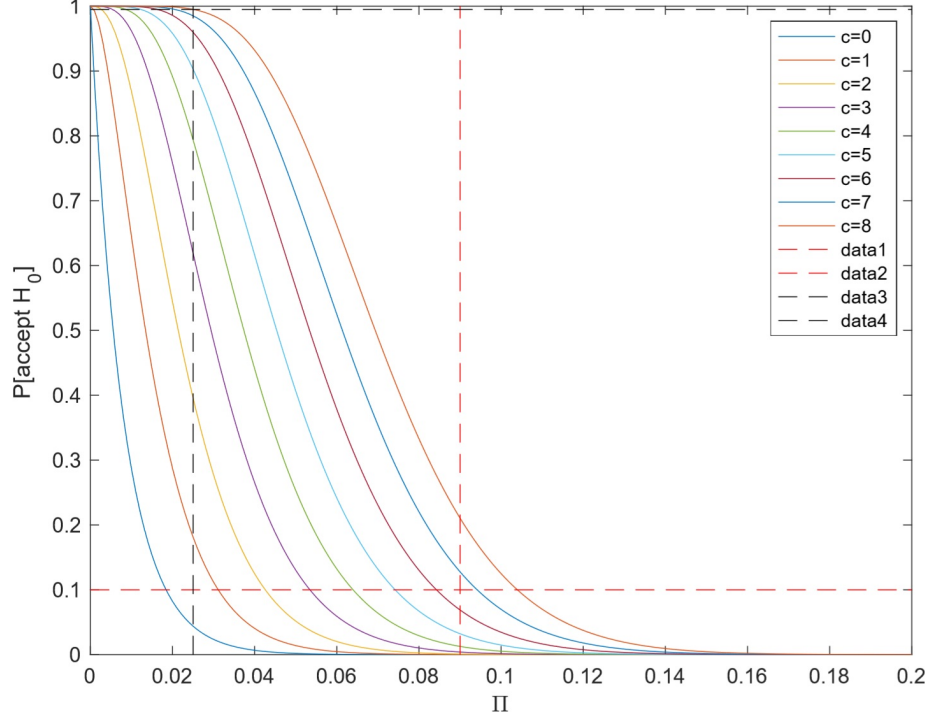


Figure 17: OC Curves when $n = 125$

For $n = 125$, we read from the plot that α requires $c > 7$ and β requires $c < 7$. Clearly it's not possible so we seek for middle solution. We found that the customer's risk is unreasonably high for larger c , and the producer's risk gets larger for smaller c . Therefore, we choose a middle solution $c = 7$.

However, there is still some deviations from what we expected. Detailed discussion and modelling we be included in Section 6, oriented to question iv).

5 Non-central T Distribution

5.1 Objectives

The objective of this section is to look into the non-central T distribution where T -test involves. In this section, we will discuss some properties of this distribution, and applied it to the derivation of OC Curves and shortfall probability.

5.2 Definition and Notations

Operating Characteristic Curve (OC Curve)[3] With the help of OC Curves, people can evaluate non-normal distributions, where finding the integral is impractical.

In a hypothesis test with null hypothesis H_0 , a single *OC Curve* is a plot of $P[\text{accepting } H_0]$, the probability of accepting(failing to reject) H_0 , versus θ , some population parameter. The single curves each correspond a particular choice of test parameters.

T-Test[3] Let X_1, \dots, X_n be a random sample of size n from a normal distribution with sample mean \bar{X} and sample variance S^2 . The normal distribution has unknown standard deviation σ , unknown population mean μ and we set up μ_0 as a null value of μ , and denote the difference

$$\delta := \mu - \mu_0.$$

In our discussion, the OC Curves for T-Test feature an abscissa d defined by

$$d := \frac{\mu - \mu_0}{\sigma} = \frac{\delta}{\sigma}$$

Non-central T-Distribution[4] The *non-central T-Distribution* is a generalization of *Student's T-Distribution*, by adding a non-centrality parameter ν . The following definition is based on [4], retrieved from Wikipedia.

Let Z and V_γ be independent random variables, where Z follows *Standard Normal Distribution* and V_γ follows *Chi-Squared Distribution*, with γ degrees of freedom. Given $\nu \neq 0 \in \mathbb{R}$, the random variable:

$$T_{\gamma, \nu} = \frac{Z + \nu}{\sqrt{V_\gamma/\gamma}}$$

is said to follow a Non-central T-Distribution with γ degrees of freedom and non-centrality parameter ν .

The direct formula of probability density function(PDF) and cumulative distribution function(CDF) of $T_{\gamma, \nu}$ are complicated and beyond discussion. Hence for short, we denote the PDF and CDF of $T_{\gamma, \nu}$ by $f_{\gamma, \nu}$ and $F_{\gamma, \nu}$.

5.3 Derivation of the OC Curve

Theorem The random variable

$$T_{n-1} = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

follows a non-central T-distribution with $n - 1$ degrees of freedom and non-centrality parameter $d\sqrt{n}$.

Inspired by Theorem 2.3.9 in [3], we basically want to construct a standard normal random variable and a chi-squared random variable in $T_{n-1} = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$.

The proof of this theorem is as follows.

$$\begin{aligned}
T_{n-1} &= \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \\
&= \frac{(\bar{X} - \mu) + (\mu - \mu_0)}{S/\sqrt{n}} \\
&= \frac{(\bar{X} - \mu) + \delta}{S/\sqrt{n}} \\
&= \frac{\frac{(\bar{X} - \mu) + \delta}{\sigma/\sqrt{n}}}{\frac{S/\sqrt{n}}{\sigma/\sqrt{n}}} \\
&= \frac{\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} + \frac{\delta}{\sigma/\sqrt{n}}}{\sqrt{[(n-1)S^2/\sigma^2]/(n-1)}} \\
&= \frac{Z + \nu}{\sqrt{\chi_{n-1}^2/(n-1)}}
\end{aligned}$$

We know that $Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$ follows a standard normal distribution, and $\chi_{n-1}^2 = \frac{(n-1)S^2}{\sigma^2}$ follows a chi-squared distribution with $n-1$ degrees of freedom.

Hence as is defined in Section 4.3, we conclude that T_{n-1} is a non-central T distribution with $n-1$ degrees of freedom and the non-centrality parameter is given by

$$\begin{aligned}
\nu &= \frac{\delta}{\sigma/\sqrt{n}} \\
&= \frac{\delta \cdot d}{\sigma/\sqrt{n}} \\
&= d\sqrt{n}
\end{aligned}$$

The proof is done.

The formula of the OC Curves is in essence a function of $P[\text{accepting } H_0]$ versus $\{d, \alpha, n\}$. In the T -Test, the null hypotheses H_0 can be either:

$$H_0 : \mu = \mu_0$$

$$H_0 : \mu \geq \mu_0$$

$$H_0 : \mu \leq \mu_0$$

The following will discuss them one-by-one.

Derivation of OC Curve When $H_0 : \mu = \mu_0$ When we set

$$H_0 : \mu = \mu_0,$$

we fix the critical region by $|t| > t_{\alpha/2, n-1}$ and reject H_0 at significance level α if

$$|T_{n-1}| > t_{\alpha/2, n-1}$$

We can then estimate the probability:

$$\begin{aligned} P[\text{accepting } H_0] &= P[|T_{n-1}| \leq t_{\alpha/2, n-1}] \\ &= P[-t_{\alpha/2, n-1} \leq T_{n-1} \leq t_{\alpha/2, n-1}] \end{aligned}$$

As is proved in the previous theorem, $T_{n-1} = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$ follows non-central T-distribution with $n - 1$ degrees of freedom and non-centrality parameter $d\sqrt{n}$. Hence we can represent the probability through integral estimation:

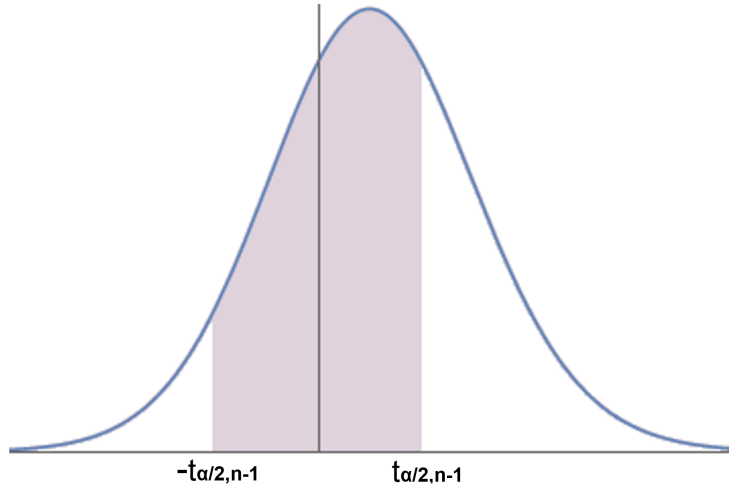


Figure 18: Two-sided Integral Estimation of a Non-Central T-Distribution

The probability is represented by the shadowed area and can be calculated by integral.

$$\begin{aligned} P[\text{accepting } H_0] &= P[-t_{\alpha/2, n-1} \leq T_{n-1} \leq t_{\alpha/2, n-1}] \\ &= \int_{-t_{\alpha/2, n-1}}^{t_{\alpha/2, n-1}} f_{n-1, d\sqrt{n}}(x) dx \\ &= F_{n-1, d\sqrt{n}}(x) \Big|_{-t_{\alpha/2, n-1}}^{t_{\alpha/2, n-1}} \\ &= F_{n-1, d\sqrt{n}}(t_{\alpha/2, n-1}) - F_{n-1, d\sqrt{n}}(-t_{\alpha/2, n-1}) \end{aligned}$$

Through the above process, we successfully derived a formula for the OC Curve of the T-test in terms of the cumulative distribution function of the non-central T distribution for $H_0 : \mu = \mu_0$:

$$P[\text{accepting } H_0] = F_{n-1, d\sqrt{n}}(t_{\alpha/2, n-1}) - F_{n-1, d\sqrt{n}}(-t_{\alpha/2, n-1})$$

Derivation of OC Curve When $H_0 : \mu \leq \mu_0$ When we set

$$H_0 : \mu \leq \mu_0,$$

we fix the critical region by $t > t_{\alpha/2, n-1}$ and reject H_0 at significance level α if

$$T_{n-1} > t_{\alpha/2, n-1}$$

Again, we can then estimate the probability

$$P[\text{accepting } H_0] = P[T_{n-1} \leq t_{\alpha/2, n-1}]$$

through integral estimation:

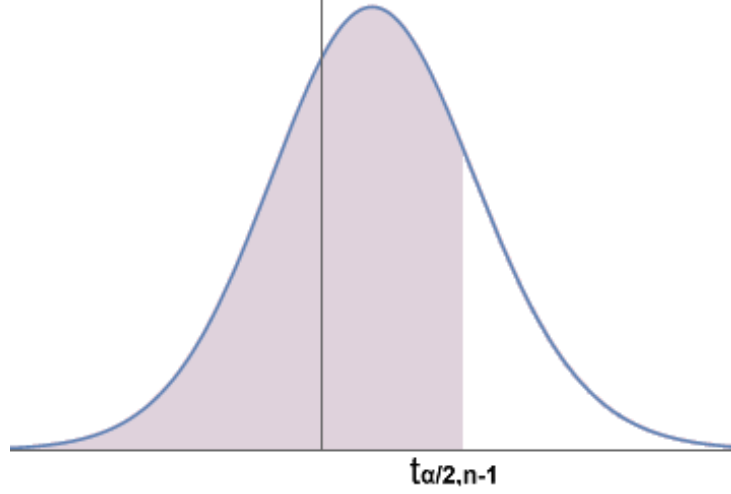


Figure 19: Upper Bounded One-sided Integral Estimation of a Non-Central T-Distribution

The probability is represented by the shadowed area and can be calculated by integral.

$$\begin{aligned} P[\text{accepting } H_0] &= P[T_{n-1} \leq t_{\alpha/2, n-1}] \\ &= \int_{-\infty}^{t_{\alpha/2, n-1}} f_{n-1, d\sqrt{n}}(x) dx \\ &= F_{n-1, d\sqrt{n}}(t_{\alpha/2, n-1}) \end{aligned}$$

Hence the formula for the OC Curve of the T-test in terms of the cumulative distribution function of the non-central T distribution for $H_0 : \mu \leq \mu_0$:

$$P[\text{accepting } H_0] = F_{n-1, d\sqrt{n}}(t_{\alpha/2, n-1})$$

Derivation of OC Curve When $H_0 : \mu \geq \mu_0$ When we set

$$H_0 : \mu \geq \mu_0,$$

we fix the critical region by $t < t_{\alpha/2, n-1}$ and reject H_0 at significance level α if

$$T_{n-1} < t_{\alpha/2, n-1}$$

Again, we can then estimate the probability

$$P[\text{accepting } H_0] = P[T_{n-1} \geq t_{\alpha/2, n-1}]$$

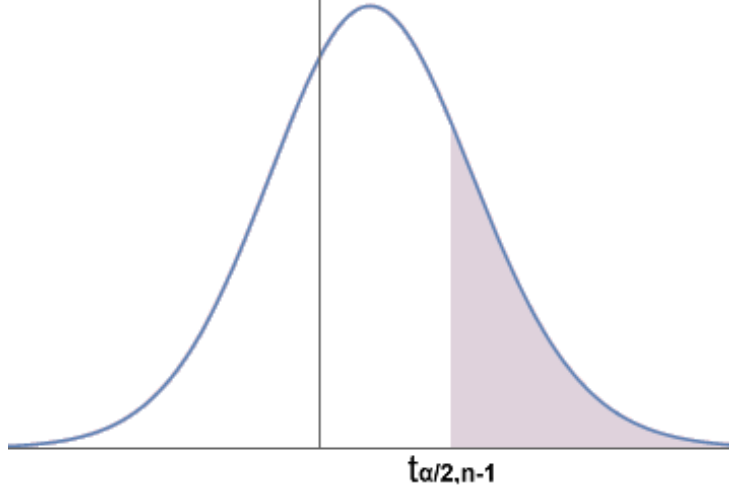


Figure 20: Lower Bounded One-sided Integral Estimation of a Non-Central T-Distribution

through integral estimation:

The probability is represented by the shadowed area and can be calculated by integral.

$$\begin{aligned}
 P[\text{accepting } H_0] &= P[T_{n-1} \geq t_{\alpha/2, n-1}] \\
 &= \int_{t_{\alpha/2, n-1}}^{\infty} f_{n-1, d\sqrt{n}}(x) dx \\
 &= 1 - F_{n-1, d\sqrt{n}}(t_{\alpha/2, n-1})
 \end{aligned}$$

Hence the formula for the OC Curve of the T-test in terms of the cumulative distribution function of the non-central T distribution for $H_0 : \mu \geq \mu_0$:

$$P[\text{accepting } H_0] = 1 - F_{n-1, d\sqrt{n}}(t_{\alpha/2, n-1})$$

5.4 The OC Curves of the Test in Table 4

We can see from Part 4.3.2 that data from Table 4 fit the third condition in T-test, with null hypothesis

$$H_0 : \mu = Q_n$$

The vertical coordinates of OC curve is $P[\text{accepting } H_0]$ and it is related to n, d, α . n and α are testing variables and d is $\frac{\mu - \mu_0}{\sigma}$.

The one-sided confidence coefficient given by Table 4 is 99.5%, so α is 0.005. We take n coming from the second column of Table 4 (10,13,50,80,125) as n in OC curve. Then we plot it using **Mathimetica**.

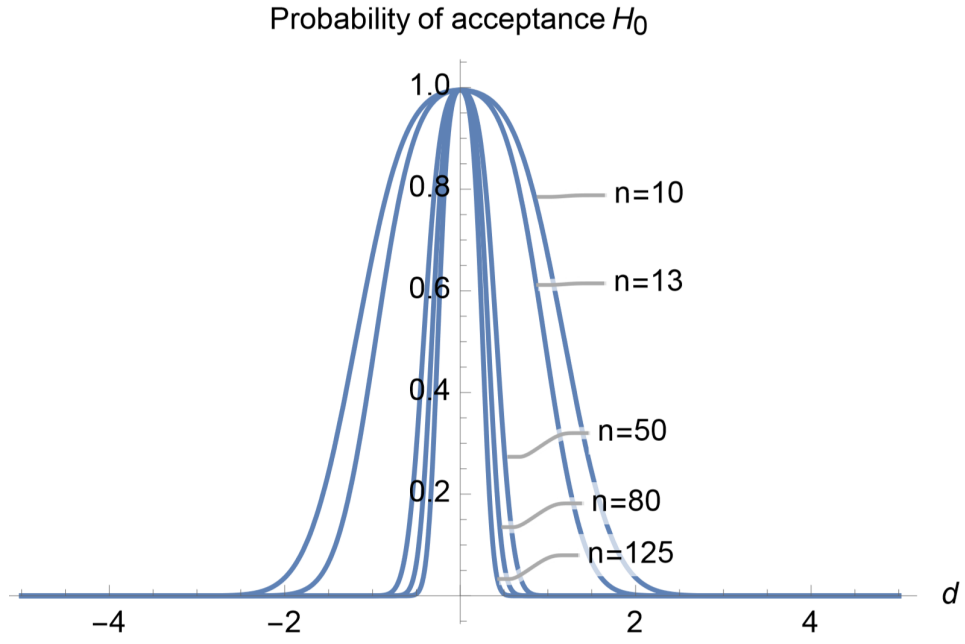


Figure 21: The OC Curves of the Test in Table 4

We alternatively give the plots for other null hypothesis for comparison.

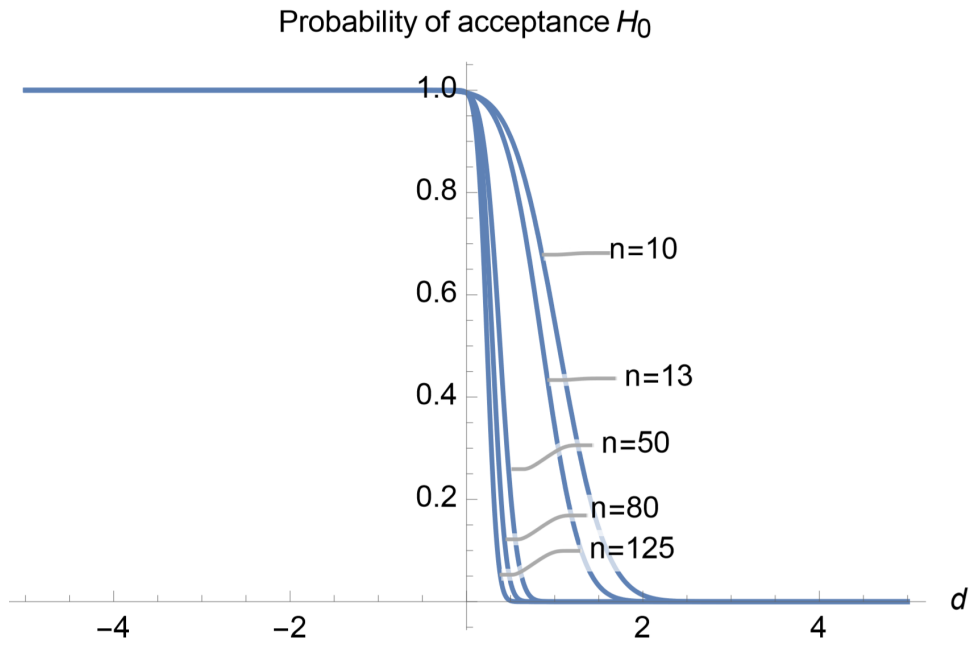


Figure 22: The OC Curves of the Test in Table 4, $H_0 : \mu \le Q_n$

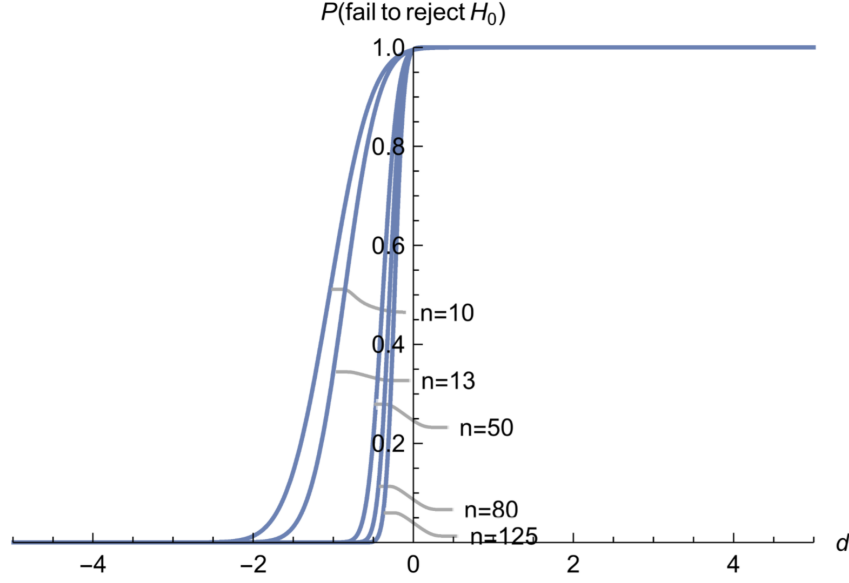


Figure 23: The OC Curves of the Test in Table 4, $H_0 : \mu \geq Q_n$

5.5 Power of the Tests

Power is defined as $1 - \beta$. Since β defines the probability that we accept H_0 even though H_1 is true, the meaning of power is that it defines the probability of successfully rejecting a false H_0 given that H_1 is true.

$$\text{power} = P[\text{reject } H_0 \mid H_1 \text{ is true}]$$

We noticed that The OC Curve is a plot of $1 - P[\text{reject } H_0]$ against some parameter that can be determined if H_1 is true, i.e. if H_1 is true, we can find the value of d and read the power from the y -axis.

Hence, β corresponds to a point on OC curve, and power is $1 - \beta$.

Observations show that:

- Let d and α be unchanged. When n increases, β decreases and power increases.
- Let α and n be unchanged. We will get $|\mu - \mu_0|$ from data and thus d . Then we will get a point on the vertical coordinates.

This identifies with the equation $n \approx \frac{(z_{\alpha/2} + z_\beta)^2 \sigma^2}{\delta^2}$ in [3], where we can see that when n increases, β decreases and power increases. As long as n is large enough, one can always reject H_0 , which decreases the customer's risk.

5.6 Probability of Detecting a Shortfall of at Least One Standard Deviation

- When $N \geq 11$, we get a sample from products, where n is the size of sample. The probability is that we detect a sample mean \bar{q} with a shortfall of at least one sample standard deviation s from the nominal net quantity Q_n .

$$\begin{aligned} P[\bar{Q} \leq Q_n - s] &= P\left[\frac{\bar{Q} - Q_n}{s/\sqrt{n}} \leq \frac{Q_n - s - Q_n}{s/\sqrt{n}}\right] \\ &= P\left[\frac{\bar{Q} - Q_n}{s/\sqrt{n}} \leq -\sqrt{n}\right] \end{aligned}$$

As is discussed, $\frac{\bar{Q} - Q_n}{s/\sqrt{n}}$ follows Non-central T Distribution.

$$\begin{aligned} P[\bar{Q} \leq Q_n - s] &= P[T_{n-1} \leq -\sqrt{n}] \\ &= F_{n-1, d\sqrt{n}}(-\sqrt{n}) \end{aligned}$$

When $d = 0$, H_0 is assumed true. This is equal to ask, in $\bar{q} \leq Q_n - \lambda s$ when the modifying factor $\lambda = 1$, find the corresponding confidence level of each n .

- When $N \leq 10$, we test all the items. Hence σ and mean μ of the products are known.

When $N \leq 10$, $n = N$ is the size of batch. The probability is that we detect a mean \bar{q} with a shortfall of at least one standard deviation σ from the nominal net quantity Q_n .

$$P[\bar{Q} \leq Q_n - \sigma] = P\left[\frac{\bar{Q} - \mu}{\sigma/\sqrt{N}} \leq \frac{Q_n - \sigma - \mu}{\sigma/\sqrt{N}}\right]$$

$\frac{\bar{Q} - \mu}{\sigma/\sqrt{N}}$ follows Standard Normal Distribution.

$$\begin{aligned} P[\bar{Q} \leq Q_n - \sigma] &= P\left[Z \leq \frac{Q_n - \sigma - \mu}{\sigma/\sqrt{N}}\right] \\ &= \Phi\left[\frac{Q_n - \sigma - \mu}{\sigma/\sqrt{N}}\right] \end{aligned}$$

6 Significance Level of Type I Error

6.1 Summary of Section 5.3.2 in [1]

Significance of T-Test Section 5.3.2.1 corresponds to the T-Test for sample mean net quantity. To summarize, for inspection lot satisfying $\mu = Q_n$, the probability of being rejected is not more than 0.5%.

To make it understandable for manufacturers, the quality supervision departments display formula $\bar{q} \geq (Q_n - \lambda s)$ where λ is shown in the third column in Table 4. The meaning is that the average sample content should be not less than the label net content minus the product of correction factor λ and the standard deviation of the sample.

Significance of Acceptance Sampling Section 5.3.2.1 corresponds to the acceptance sampling for detective rate. To summarize, for inspection lot satisfying the fact that the percentage of shortage quantitative packaged goods is not more than 2.5%, the probability of being rejected is not more than 5%.

To be more specific, the shortage quantitative packaged goods include both T_1 shortfall products and T_2 shortfall products. It can be used to calculate the allowed amounts of T_1 shortfall and T_2 shortfall, which is shown in the forth column in Table 4 in JTF 1070-2005.

6.2 Discussions of the Statements

From above discussions, the significant value α can be approximately calculated as:

$$\alpha \approx \sum_{d>c} \frac{(n\Pi_0)^d e^{(-n\Pi_0)}}{d!} \quad (3)$$

in which $\Pi_0 = 2.5\%$ and c is shown in Table 4 as 0,0,1,3,5,7. d is $\min(n, r)$ where $r = \text{floor}(N\Pi_0)$.

Both significant levels above are usually called the Type I error(α) or the producer's risk. On one hand, it guarantees the profit of producers by limiting the probability of rejection. On the other hand, it shows the maximum allowed Type T_1 and T_2 shortfall. The specific calculation is displayed in the above analysis. In the Section 4.3.1, we discuss the level of Type I error using three methods including hypergeometric distribution, approximating binomial and poisson distribution. Table 3 and 4 show the corresponding α value which is quite larger than the threshold value 5% when the sample amount is small. The tables also show a tendency of α to decrease below 5%. Some reasons can explain this abnormal Type I error:

- The 0.5% significant value limits the mean of the products and the 5% significant value limits the passing rate. The abnormal value in Table 3 and 4 can be caused by the balance of producer's risk and consumer's risk.
- Noticing when the inspection lot amount is larger than 3200, the sample amount is 125. In Table 3 and 4, the corresponding α value is below 5%. In manufacturing, the larger amount of the

products in prepackages with fixed content, the more necessary to control its cost. Rejecting too many acceptable products not only wastes materials but also increases the cost. Such as the car wheels production, the small preset α value protects the manufacturer's profit. Similarly, when the inspection lot has a rather small amount, the high level of rejection probability ensures the quality of the products. Owing to the small amount, the manufacturer can improve the process better.

- Consideration that T_2 shortfall will be taken into account for smaller Q_n . Note that in the previous sections, we have assume that T_2 shortfalls are ignorable.

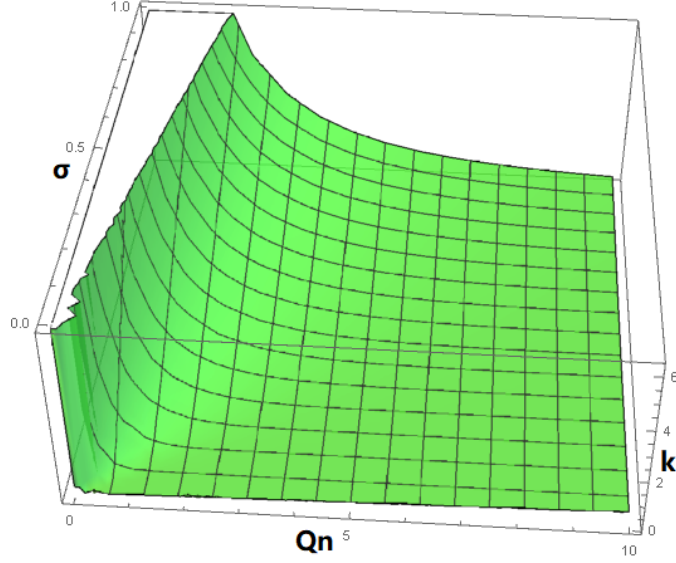


Figure 24: The Plot of k against σ and Q_n

However, for batch of products with small Q_n and large σ , this is not the case. For example, we fix $\sigma = 1$ and $\mu = Q_n = 5$, we can calculate the k value is only:

$$k = 1.29351$$

This would mean that a T_2 shortfall will be detected more possibly than a T_1 shortfall!

- Consideration of a barely acceptable Π_1 .

We assume that for larger N , we define $H_1 : \Pi > \Pi_1 = 0.12$, and plot the c versus N using Hypergeometric distribution. The result is as follows.

We found that for larger N , $c = 7$ is true.

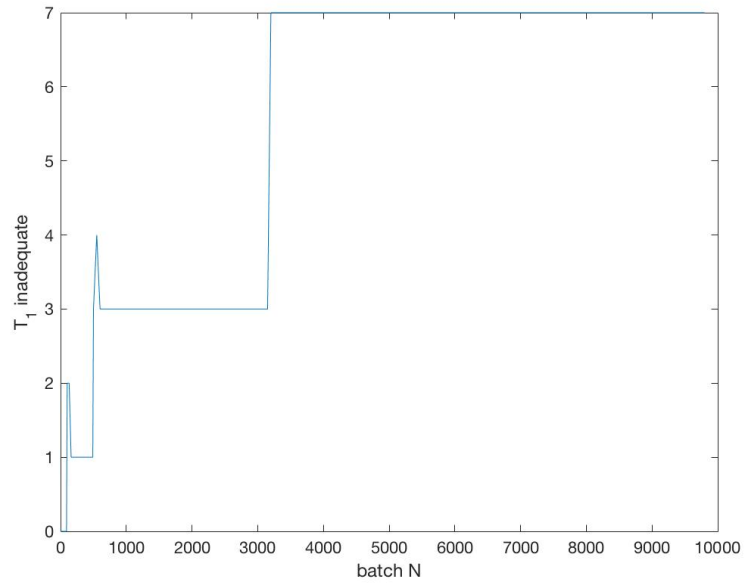


Figure 25: The Relationship between c_{\max} and N , Hypergeometric, $\Pi_1 = 0.12$

7 Tests of a Sample of Prepackaged Food

7.1 Objectives

The goal of this exercise is to test on a package of DOUBLEMINT pressed candies and verify whether the actual package contents correspond to the nominal contents according to JTF 1070-2005.

7.2 Testing Materials and Apparatus

DOUBLEMINT pressed candies, electronic weigh.

7.2.1 Testing Materials



Figure 26: DOUBLEMINT pressed candies

Net weight	12g
Production date	2018.08.15
Shelf life	12 months

Table 12: Main Labels on the Package

7.2.2 Apparatus

Electronic scales



Figure 27: Electronic Scale

Units	g
Accuracy	0.1g
Capacity	1000g
Operation Temp	10 to 30 dec C

Table 13: Main Parameters of the Scale

7.3 Procedure

7.3.1 Measurement of average tare weight

- Select two packages from the sample. Measure the difference of their net weight (R_c) and the difference of the tares (R_t).
- Choose the number of packages in the sample that we need to measure the tares(n_i) with the value R_c/R_t according to Table B.2.

- If $R_c/R_t \leq 1.00$, test the weight of every tare as the tare weight of the package. Else we test the average weight of all n_i tares as the tare weight.



Figure 28: Measurement of Net Weight



Figure 29: Measurement of Tare Weight



Figure 30: The general weight of package

7.3.2 Measurement of the actual content deviation

- Measure the general weight (GW_i) of every package.
- Calculate the actual content (q_i) of every package with the equation $q_i = GW_i - \bar{P}$.
- Calculate the content deviation of every package (D_i) with the equation $D_i = q_i - Q_n$.

7.4 Measurements and Results

We had 99 packages of DOUBLEMINT pressed candies which was the inspection lot. According to [3, Table 4], we chose 13 of them as the sample that we tested on. For convenience, we labelled them from Package 1 to 13.

7.4.1 Measurement of average tare weight

Select Package 1 and 2 and we measured:

i	Net weight(measured) [g]	Tare weight(measured) [g]
1	12.8	8.4
2	12.8	8.3

Table 14: Net and Tare Weight of the Samples

Since the measured $R_c = 0$, we chose ' $R_c/R_t \leq 0.2$ ' in Table 2. The corresponding tare sample number n_i is 13. $n_i = n$, so we would need to measure the actual tare weight for every package.

i	Tare weight(measured) [g]	i	Tare weight(measured) [g]	i	Tare weight(measured) [g]
1	8.4	2	8.3	3	8.3
4	8.3	5	8.0	6	8.3
7	8.3	8	8.3	9	8.4
10	8.6	11	8.4	12	8.3
13	8.4				

Table 15: All Measurements of Tare Weights

7.4.2 Actual content deviation

The results are as follows.

i	GW_i [g]	$q_i = GW_i - \bar{P}$ [g]	Q_n [g]	$D = q_i - Q$ [g]
1	21.2	12.8	12	0.8
2	21.1	12.8	12	0.8
3	21.1	12.8	12	0.8
4	21.1	12.8	12	0.8
5	21.2	13.2	12	1.2
6	20.9	12.6	12	0.6
7	21.1	12.8	12	0.8
8	21.2	12.9	12	0.9
9	21.2	12.8	12	0.8
10	21.4	12.8	12	0.8
11	21.2	12.8	12	0.8
12	21.4	13.1	12	1.1
13	21.2	12.8	12	0.8

Table 16: Calculation of Actual Content Deviation

We can then plot the boxplot of the obtained data.

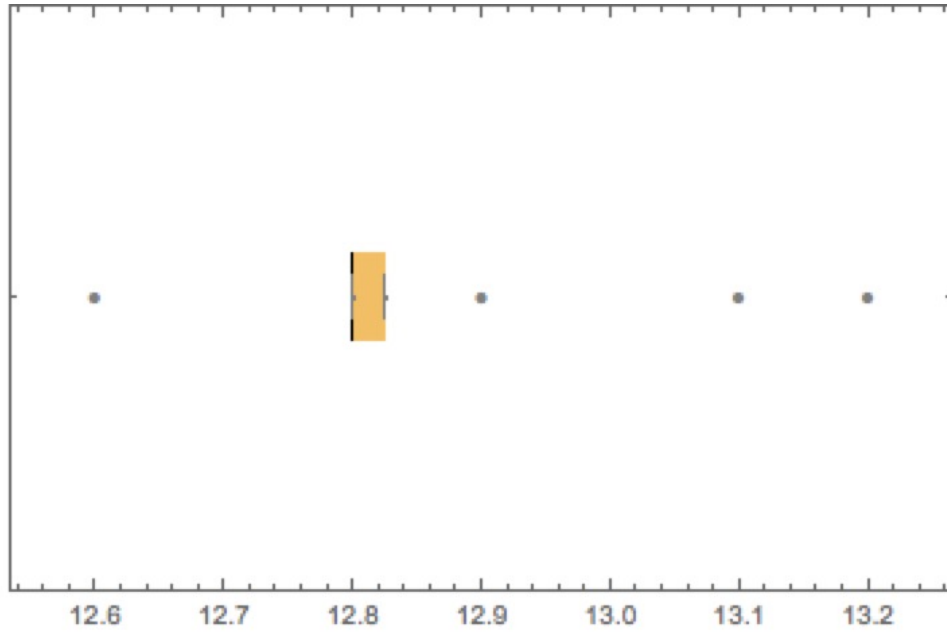


Figure 31: The boxplot of sample weight

7.5 Analysis

First we calculate the average content of samples:

$$\begin{aligned}\bar{q} &= \frac{1}{n} \sum_{i=1}^n q_i \\ &= \frac{1}{13} (12.8 * 9 + 13.2 + 13.1 + 12.6 + 12.9) [g] \\ &= 12.85 [g]\end{aligned}$$

Then we get the standard deviation of the sample actual contents:

$$\begin{aligned}s &= \sqrt{\frac{1}{n-1} \sum_{i=1}^n (q_i - \bar{q})^2} \\ &= \sqrt{\frac{1}{12} (0.05^2 \times 10 + 0.25^2 \times 2 + 0.35^2)} [g] \\ &= \sqrt{\frac{1}{12} \times 0.2725} [g] \\ &= 0.151 [g]\end{aligned}$$

In Table 4, the modifying factor $\lambda = t_{0.995} * \frac{1}{\sqrt{n}} = 0.848$ when $n = 13$ is provided, so we have:

$$Q_n - \lambda s = 12 - 0.848 \times 0.151 [g] = 11.87 [g]$$

Thus, we have clear evidence applying the testing rules for net quantities of a batch:

- The average content of samples $\bar{q} = 12.85 [g]$ is larger than the net quantity minus the actual average quantity modification value $Q_n - \lambda s = 11.87 [g]$.
- The number of type T_1 shortfalls is less than the preset acceptance number 1.
- No type T_2 shortfalls occurs for the sample weights are all larger than the net quantity.

In all, the package of DOUBLEMENT pressed candies **accords with** the measurement requirements in JTF 1070-2005.

Further, something needs to be noticed:

- The standard deviation s has a rather small value 0.151g. In Figure 13, we can see four points are near outliers and the distance between adjacent values is rather small. Thus, there is evidence of this small deviation and symmetrical feature. The larger standard deviation s is, the more necessary to increase the preset mean of normal distribution in manufacturing to reduce the consumer's risk. This undoubtedly increases the average cost of production. So a wiser way is to improve the accuracy of the production machine.
- Our sample weights are all larger than the nominal quantity. The 99.5% confidence interval ensures most of the sample net quantity is much larger than the nominal quantity.

- The amount of inadequate products in prepackages with fixed content is strictly limited. When the shortfall occurs (especially type T_1 shortfall), finding the reason for the abnormal fluctuation is on the top priority. After eliminating the interfering factor, we can do the sampling inspection again and put back on production. If the type T_2 shortfall is detected, it means the production machine has a lower accuracy of manufacturing. Then the production should be stopped immediately and manufacturer should rectify and reform the process capability.
- The confidence coefficient is 99.5% rather than 100% in this metrology inspection. There exists the Type II error(β) during the inspection. Generally, we use other methods to lower this risk such as ‘Six Sigma’ strategies [5], control chart [6] and process capability index etc.

8 Conclusion

In conclusion, the project investigate the *Rules of metrological testing for net quantities of products in prepackages with fixed content*, issued by Chinese government in 2005.

Specifically, we are interested in the mean testing and acceptance sampling methods and rules that Chinese government predetermined to measure the net quantity of a batch of products and control the type I error (rejecting an acceptable lot) and type II error (failing to reject an unacceptable lot) by verification of the significance and acceptance numbers.

We successfully conduct tests with accord to the document, in order to verify whether or not the actual quantity of DOUBLEMINT sample corresponds to its labelled contents. We finally conclude that the package of DOUBLEMENT pressed candies **accords with** the measurement requirements in JTF 1070-2005.

9 References

9.1 Works Cited

- [1] Rules of metrological testing for net quantities of products in prepackages with fixed content. *Technical Report JJF 1070-2005*, China Metrology Publishing House, 2005. <http://www.ahhz.gov.cn/CountyOpennessContent/show/1078758.html> [Online; accessed 22-March-2019].
- [2] Requiem for the Truth-in-Packaging Bill? *Journal of Marketing*, Vol.30, No.2 (Apr 1966), pp. 1-5, JSTOR. <https://www.jstor.org/stable/1249054> [Online; accessed 3-April-2019].
- [3] H. Hohberger. "ve401_main.pdf" (2018). UMJI-SJTU, Shanghai. [Online; accessed 25-Feb-2019].
- [4] Wikipedia. Noncentral t-distribution, *wikipedia*, the free encyclopedia, 2014. https://en.wikipedia.org/wiki/Noncentral_t-distribution [Online; accessed 10-April-2014].
- [5] Wikipedia. Six Sigma, *wikipedia*, the free encyclopedia, 2014. https://en.wikipedia.org/wiki/Six_Sigma [Online; accessed 10-April-2014].
- [6] Wikipedia. Control chart, *wikipedia*, the free encyclopedia, 2014. https://en.wikipedia.org/wiki/Control_chart [Online; accessed 10-April-2014].

9.2 Figures

- [1] https://en.wikipedia.org/wiki/Central_limit_theorem#/media/File:Central_Limit_Theorem.png

10 Appendix

10.1 Mathematica and Matlab Codes

10.1 Figure for Student's T-Distributions in Section 3.4

The Mathematica codes for Figure t is as follows.

```
1      Plot[PDF[StudentTDistribution[25], k]
2      // Evaluate, {k, -6, 8},
3      Ticks -> Null,
4      PlotRange -> {{-4, 4}, {0, 0.4}},
5      Filling -> Axis]
```

10.2 Figure for T_1 and T_2 Shortfall Analysis in Section 4.2

The Mathematica codes for Figure t is as follows.

```
1      Clear["`*"]
2      T = Piecewise[{{0.09*Qn, Qn <= 50},
3      {4.5, 50 < Qn <= 100},
4      {0.045*Qn, 100 < Qn <= 200},
5      {9, 200 < Qn <= 300},
6      {0.03*Qn, 300 < Qn <= 500},
7      {15, 500 < Qn <= 1000},
8      {0.015*Qn, 1000 < Qn <= 10000},
9      {150, 10000 < Qn <= 15000},
10     {0.01*Qn, 15000 < Qn <= 50000}}];
11     mu = Qn;
12     PA = 1 - CDF[NormalDistribution[mu, sigma], Qn - T];
13     PT1 = CDF[NormalDistribution[mu, sigma], Qn - T] - \
14     CDF[NormalDistribution[mu, sigma], Qn - 2*T];
15     PT2 = CDF[NormalDistribution[mu, sigma], Qn - 2*T];
16     k = PT2/PT1;
17     Plot3D[k, {Qn, 0, 10}, {sigma, 0, 1},
18     PlotStyle -> Directive[RGBColor[0.32, 1., 0.28], Opacity[0.814]]]
```

10.3 Figure and Calculations for Type I Significance in Section 4.3.1

The Matlab codes for Figure t is as follows.

```

1  clear;clc
2
3  %poisson
4      ret=[];
5
6  r n=[10,13,50,80,125]
7  cn(n);
8  m=0;
9  r d=c:1:n
10 m=sum+(n*0.025)^d*exp(-n*0.025)/factorial(d);
11 d
12 t=[ret,sum];
13 d
14 }
15
16 %binomial
17     ret=[];
18     for n=[10,13,50,80,125]
19         c=cn(n);
20         sum=0;
21         for d=c:1:n
22             sum=sum+nchoosek(n,d)*0.025^d*0.975^(n-d);
23         end
24         ret=[ret,sum];
25     end
26 }
27
28 hypergeometric
29 et=[];
30 N=[0:1:10,11:2:50,100:10:500,500:40:5000];
31 or N=[0:1:10,11:2:50,100:10:500,500:40:5000]
32 n=nN(N);
33 r=round(0.025*N);
34 c=cn(n);

```

```

35 sum=0;
36         for d=c:1:floor(min(n,r))
37             sum=sum+nchoosek(r,d)*nchoosek(N-r,n-d)/nchoosek(N,n);
38         end
39         ret=[ret,sum];
40     nd
41     plot(NN,ret,'-');
42
43
44     unction c=cn(n)
45     =0*(n==10)+1*(n==13)+3*(n==50)+5*(n==80)+7*(n==125);
46     nd
47
48     unction n=nN(N)
49     =N.*(N>=1&N<=10)+10*(N>=11&N<=50)+13*(N<=99&N>=51)+50*(N<=500 &      N>=100)+80*(N<=3200&N>=
50     nd
51
52     unction c=cN(N)
53     =cn(nN(N));
54     nd

```

10.4 Figures for Type II Significance in Section 4.3.2

The Matlab codes for Figure x is as follows.

```

1 lear;clc
2
3 poisson \betaa
4 for n=[10,13,50,80,125]
5     ret=[];
6     c=cn(n);
7     sum=0;
8     p=0:0.001:1;
9     for d=0:1:c
10         sum=sum+(n.*p).^d.*exp(-n.*p)./factorial(d);

```

```

11     end
12     ret=[ret,sum];
13     plot(p,ret);
14     hold on;
15 end
16 legend('n=10','n=13','n=50','n=80','n=125');
17 xlabel('\Pi');
18 label('\beta');
19
20 binomial \beta
21 for n=[10,13,50,80,125]
22     ret=[];
23 c=cn(n);
24 sum=0;
25 p=0:0.001:1;
26     for d=0:1:c
27         sum=sum+nchoosek(n,d).*p.^d.*(1-p).^(n-d);
28     end
29     ret=[ret,sum];
30     plot(p,ret);
31     hold on;
32 end
33 legend('n=10','n=13','n=50','n=80','n=125');
34 xlabel('\Pi');
35 label('\beta');
36
37 unction c=cn(n)
38 =0*(n==10)+1*(n==13)+3*(n==50)+5*(n==80)+7*(n==125);
39 nd
40
41 unction n=nN(N)
42 =N.*(N>=1&N<=10)+10*(N>=11&N<=50)+13*(N<=99&N>=51)+50*(N<=500
& N>=100)+80*(N<=3200&N>=
43 nd
44
45 unction c=cN(N)

```

```
46 =cn(nN(N));
```

```
47 nd
```

10.5 Figures for c v.s. N in Section 4.4.1 and 4.4.2

The Matlab codes for Figure x is as follows.

```
1 lear;clc
2 et=[];
3
4 Hypergeometric: \alpha
5 or N=[1:1:10,11:3:50,100:10:500,500:400:5000]
6
7         n=nN(N);
8
9 =round(0.025*N);
10
11 lag=0;
12 _min=0;
13 or c=0:1:N
14     sum=0;
15     for d=(c+1):1:floor(min(n,r))
16         sum=sum+nchoosek(r,d)*nchoosek(N-r,n-d)/nchoosek(N,n);
17     end
18     if sum<=0.05
19         c_min=c;
20         break
21     end
22 nd
23 et=[ret,c_min];
24 nd
25
26 Hypergeometric: \beta
27 or     N=[0:1:10,11:2:50,51:10:99,101:30:500,501:50:3200,3201:600:10000]
28 =nN(N);
```

```

29 =round(0.09*N);
30 _max=0;
31 or c=10:-1:0
32         sum=0;
33         for d=max((n+r-N),0):1:min(c,min(n,r))
34             sum=sum+nchoosek(r,d)*nchoosek(N-r,n-d)/nchoosek(N,n);
35         end
36         if sum<=0.1
37             c_max=c;
38             break
39         end
40 nd
41 et=[ret,c_max];
42 nd
43 N=[0:1:10,11:2:50,51:10:99,101:30:500,501:50:3200,3201:600:10000];
44 lot(NN,ret,'-');
45
46 binomial distribution \alpha
47 or N=[1:1:10,11:3:50,100:10:500,500:400:5000]
48 =nN(N);
49 =round(0.025*N);
50 lag=0;
51 _min=0;
52 or c=0:1:r
53         sum=0;
54         for d=0:1:c
55             sum=sum+nchoosek(n,d)*0.025^d*0.975^(n-d);
56         end
57         if 1-sum<=0.05
58             c_min=c;
59             break
60         end
61 nd
62 et=[ret,c_min];
63 nd

```

```

64 N=[1:1:10,11:3:50,100:10:500,500:400:5000];
65 lot(NN,ret,'-');
66
67 binomial: \beta
68 or N=[1:1:10,11:2:50,100:10:500,500:40:5000]
69 =nN(N);
70 =round(0.09*N);
71 _max=0;
72 or c=10:-1:0
73     sum=0;
74     for d=0:1:min(c,n)
75         sum=sum+nchoosek(n,d)*0.09^d*0.91^(n-d);
76     end
77     if sum<=0.1
78         c_max=c;
79         break
80     end
81 nd
82 et=[ret,c_max];
83 nd
84 N=[1:1:10,11:2:50,100:10:500,500:40:5000];
85 lot(NN,ret,'-');
86
87 oisson: \alpha
88 or N=[1:1:10,11:2:50,100:1:500,500:1:5000]
89     n=nN(N);
90 lag=0;
91 _min=0;
92 =0.025*N;
93 or c=0:1:N
94     sum=0;
95     for d=(c+1):1:floor(min(n,r))
96         sum=sum+(0.025*n)^d*exp(-0.025*n)/factorial(d);
97     end
98     if sum<=0.05

```

```

99         c_min=c;
100         break
101     end
102 nd
103 et=[ret,c_min];
104 nd
105
106 N=[1:1:10,11:2:50,100:1:500,500:1:5000];
107 lot(NN,ret,'-');
108
109 poisson: \beta
110 or N=[1:1:10,11:2:50,100:1:500,500:1:5000]
111 =nN(N);
112 =round(0.09*N);
113 _max=0;
114 or c=10:-1:0
115     sum=0;
116     for d=0:1:c
117         sum=sum+(0.09*n)^d*exp(-0.09*n)/factorial(d);
118     end
119     if sum<=0.1
120         c_max=c;
121         break
122     end
123 nd
124 et=[ret,c_max];
125 nd
126 N=[1:1:10,11:2:50,100:1:500,500:1:5000];
127 lot(NN,ret,'-');
128
129 unction n=nN(N)
130 =N.*(N>=1&N<=10)+10*(N>=11&N<=50)+13*(N<=99&N>=51)+50*(N<=500
& N>=100)+80*(N<=3200&N>=
131 nd

```

10.6 for OC Curves of Π and c in Section 4.4.3

The Matlab codes for Figure x is as follows.

```
1 lear;clc
2 =125;
3 or c=0:1:8
4     x=0:0.001:0.2;
5     sum=0;
6     for d=0:1:c
7         sum=sum+(n.*x).^d.*exp(-n.*x)./factorial(d);
8     end
9     y=sum;
10    plot(x,y,'-');
11
12    hold on;
13 nd
14 legend('c=0','c=1','c=2','c=3','c=4','c=5','c=6','c=7','c=8');
15
16 1=line([0,0.2],[0.1,0.1],'linestyle','--');
17 1.Color='red';
18 2=line([0.09,0.09],[0,1],'linestyle','--');
19 2.Color='red';
20
21 3=line([0,0.2],[0.995,0.995],'linestyle','--');
22 3.Color='black';
23 4=line([0.025,0.025],[0,1],'linestyle','--');
24 4.Color='black';
25 label('\Pi');
26 label('P[accept H_0]');
```

10.7 for OC Curves of d and n in Section 5.4

The Mathematica codes for Figure x is as follows.

```

1 [CDF[NoncentralStudentTDistribution[\[Nu], 0], t], 50]
2
3
4 *Two-sided mu=mu0*****)
5 1 = Plot[-CDF[NoncentralStudentTDistribution[9, Sqrt[9] d],
6     x = -h[0.0025, 9]] +
7     CDF[NoncentralStudentTDistribution[9, Sqrt[9] d],
8     x = h[0.0025, 9]], {d, 0, 5},
9 PlotLabels -> {"n=10", Placed[Automatic, Above]}]
10
11 2 = Plot[-CDF[NoncentralStudentTDistribution[12, Sqrt[12]*d],
12     x = -h[0.0025, 12]] +
13     CDF[NoncentralStudentTDistribution[12, Sqrt[12] d],
14     x = h[0.0025, 12]], {d, 0, 5},
15 PlotLabels -> {"n=13", Placed[Automatic, Above]}]
16
17 3 = Plot[-CDF[NoncentralStudentTDistribution[49, Sqrt[49]*d],
18     x = -h[0.0025, 49]] +
19     CDF[NoncentralStudentTDistribution[49, Sqrt[49] d],
20     x = h[0.0025, 49]], {d, 0, 5},
21 PlotLabels -> {"n=50", Placed[Automatic, Above]},
22 PlotRange -> {{0, 5}, {0, 1}}]
23
24 4 = Plot[-CDF[NoncentralStudentTDistribution[79, Sqrt[79]*d],
25     x = -h[0.0025, 79]] +
26     CDF[NoncentralStudentTDistribution[79, Sqrt[79] d],
27     x = h[0.0025, 79]], {d, 0, 5},
28 PlotLabels -> {"n=80", Placed[Automatic, Above]},
29 PlotRange -> {{0, 5}, {0, 1}}]
30 5 = Plot[-CDF[NoncentralStudentTDistribution[124, Sqrt[124]*d],
31     x = -h[0.0025, 124]] +
32     CDF[NoncentralStudentTDistribution[124, Sqrt[124] d],
33     x = h[0.0025, 124]], {d, 0, 5},
34 PlotLabels -> {"n=125", Placed[Automatic, Above]},

```

```

35 PlotRange -> {{0, 5}, {0, 1}}]
36
37 1 = Show[w1, w2, w3, w4, w5, PlotRange -> {{0, 5}, {0, 1}}]
38
39
40 *One-sided mu<mu0*****)
41 1 = Plot[-CDF[NoncentralStudentTDistribution[9, Sqrt[9]*d],
42   x = -\[Infinity]] +
43   CDF[NoncentralStudentTDistribution[9, Sqrt[9] d],
44   x = h[0.005, 9]], {d, 0, 5},
45 PlotLabels -> {"n=13", Placed[Automatic, Above]]]
46
47 2 = Plot[-CDF[NoncentralStudentTDistribution[12, Sqrt[12]*d],
48   x = -\[Infinity]] +
49   CDF[NoncentralStudentTDistribution[12, Sqrt[12] d],
50   x = h[0.005, 12]], {d, 0, 5},
51 PlotLabels -> {"n=13", Placed[Automatic, Above]]]
52
53 3 = Plot[-CDF[NoncentralStudentTDistribution[49, Sqrt[49]*d],
54   x = -\[Infinity]] +
55   CDF[NoncentralStudentTDistribution[49, Sqrt[49] d],
56   x = h[0.005, 49]], {d, 0, 5},
57 PlotLabels -> {"n=50", Placed[Automatic, Above]}],
58 PlotRange -> {{0, 5}, {0, 1}}]
59
60 4 = Plot[-CDF[NoncentralStudentTDistribution[79, Sqrt[79]*d],
61   x = -\[Infinity]] +
62   CDF[NoncentralStudentTDistribution[79, Sqrt[79] d],
63   x = h[0.005, 79]], {d, 0, 5},
64 PlotLabels -> {"n=80", Placed[Automatic, Above]}],
65 PlotRange -> {{0, 5}, {0, 1}}]
66 5 = Plot[-CDF[NoncentralStudentTDistribution[124, Sqrt[124]*d],
67   x = -\[Infinity]] +
68   CDF[NoncentralStudentTDistribution[124, Sqrt[124] d],
69   x = h[0.005, 124]], {d, 0, 5},

```

```

70 PlotLabels -> {"n=125", Placed[Automatic, Above]},
71 PlotRange -> {{0, 5}, {0, 1}}]
72
73 2 = Show[e1, e2, e3, e4, e5, PlotRange -> {{0, 5}, {0, 1}}]
74
75
76 *One-sided mu>mu0*****)
77 1 = Plot[-CDF[NoncentralStudentTDistribution[9, Sqrt[9]*d],
78   x = -h[0.005, 9]] +
79   CDF[NoncentralStudentTDistribution[9, Sqrt[9] d],
80   x = \[Infinity]], {d, -5, 5},
81 PlotLabels -> {"n=10", Placed[ Automatic, None]},
82 PlotRange -> {{-5, 5}, {0, 1}}]
83
84 2 = Plot[-CDF[NoncentralStudentTDistribution[12, Sqrt[12]*d],
85   x = -h[0.005, 12]] +
86   CDF[NoncentralStudentTDistribution[12, Sqrt[12] d],
87   x = \[Infinity]], {d, -5, 5},
88 PlotLabels -> {"n=13", Placed[Automatic, None], Style["down", 30]},
89 PlotRange -> {{-5, 5}, {0, 1}}]
90
91 3 = Plot[-CDF[NoncentralStudentTDistribution[49, Sqrt[49]*d],
92   x = -h[0.005, 49]] +
93   CDF[NoncentralStudentTDistribution[49, Sqrt[49] d],
94   x = \[Infinity]], {d, -5, 5},
95 PlotLabels -> {"n=50", Placed[Automatic, None]},
96 PlotRange -> {{-5, 5}, {0, 1}}]
97
98 4 = Plot[-CDF[NoncentralStudentTDistribution[79, Sqrt[79]*d],
99   x = -h[0.005, 79]] +
100   CDF[NoncentralStudentTDistribution[79, Sqrt[79] d],
101   x = \[Infinity]], {d, -5, 5},
102 PlotLabels -> {"n=80", Placed[Automatic, None]},
103 PlotRange -> {{-5, 5}, {0, 1}}]
104 5 = Plot[-CDF[NoncentralStudentTDistribution[124, Sqrt[124]*d],

```

```

105      x = -h[0.005, 124]] +
106      CDF[NoncentralStudentTDistribution[124, Sqrt[124] d],
107      x = \[Infinity]], {d, -5, 5},
108      PlotLabels -> {"n=125", Placed[Automatic, None]},
109      PlotRange -> {{-5, 5}, {0, 1}}]
110
111 3 = Show[d1, d2, d3, d4, d5, PlotRange -> {{-5, 5}, {0, 1}}]

```

10.8 Figures for Non-central T-Distributions in Section 5.3

The Mathematica codes for Figure x is as follows.

```

1  Plot[PDF[NoncentralStudentTDistribution[25, 2], k]
2      // Evaluate, {k, -6, 8},
3      Ticks -> Null,
4      PlotRange -> {{-2, 6}, {0, 0.4}},
5      Filling -> Axis]

```
