Project Manual: A Brief Introduction Toward the Statistics Behind The Food Inspection Document JJF1070-2005

April 8, 2019

Zhu Jing ID:517370910135
Shen Yueyang ID:517370910226
Yang Zhe ID:517021911127
Ma Siyin ID:517370910003
Gao Shiyue ID: 515370910191

Abstract

Throughout this project, we studied some of the sampling rules established in the food packaging inspection process. We delved into different aspects of the single attribute sampling schemes in document JJF1070-2005. We first explored how the equation $\overline{q} \geq$(Q_n-λ s)in Table 4 is derived, and we also discussed the relationship between the equation and the hypothesis test. The statistics behind the fifth column of Table 4 is examined, where we separate our discussion into two major scenarios: $\frac{n}{N} \leq 0.1$ where approximations can be used, and $\frac{n}{N} \geq 0.1$ where we can only use hypergeometric distribution. In order to derive an OC curve for the T-test with $\overline{q} \neq Q_n$, we use the non-central T-distribution. Some comments about the OC curve were given. To model real life situations, we also conduct a sampling process using a sample of chocolate wafers according to the procedure described in document JJF1070-2005. We found that the net quantities of all our samples are larger than the nominal quantity and the quantity distribution follows a non-central T-distribution. We wish our report will provide some useful insights in terms of understanding the mathematical and statistical details behind those sampling rules. We tried to fill in as much detail into the explanation and hopefully, our logic is easy to follow.

Contents

1 Introduction

As a rigorous document to provide guidance in terms of food production and metrological inspection, JJF1070-2005 has standardized the Chinese food production process in many ways since been put into effect in 2006[1]. In general, the rules established and the technical details in JJF1070-2005 coincide with the international standard[13]. However, this document also filled in some gaps in terms of the sampling rules regarding small lot size[13]. The purpose of this project is to discuss the statistical background of certain sections of this document and apply some of the knowledge to perform some real-life testing. Specifically, we would delve into the sampling rules set forth in Section 4.3.2 in JJF1070-2005.

1.1 Testing rules in 4.3.2

In principle, for any quantitative packaged goods inspected the average net quantity should be greater or equal to the nominal quantity.

When using a single attribute sampling plan to decide whether one would reject a lot or not, one should obey the following rules:

- For lots with size 11-50, 51-99, 100-500, 501-3200, ≥3200 their corresponding sample sizes are 10, 13, 50, 80, 125, respectively, whereas, for the lot with a size smaller than 10, the sample size should be equal to the lot size.
- For samples with size 10, 13, 50, 80, 125, the sample average net quantity should be greater or equal to nominal quantity minus a correction factor(λs), where s is the sample standard deviation, and λ takes on values 1.028, 0.848, 0.379, 0.295, 0.234 respectively.
- For a sample size smaller than 10, the number of packages which are short of the stated quantity by a defined amount(TNE-"Tolerable Negative Error" or T1 shortage) should be less than 0[14](One do not tolerate any such packages).
- For samples with size 13, 50, 80, 125, the number of packages which are short of the stated quantity by a defined amount(TNE-"Tolerable Negative Error" or T1 shortage) should be less than 1, 3, 5, 7 respectively[14].
- For any sample size less than or equal 10, one should reject a lot if there is any package in the sample that indicates T2 deficiency.

2 The Confidence Interval for the Mean Package Contents and the Hypothesis Test

The equation $\bar{q} \geq (Q_n - \lambda s)$ can be rewritten as follows.

$$\bar{q} \geq (Q_n - \lambda s) \qquad (*)$$

$\Leftrightarrow$

$$\bar{q} \geq (Q_n - t_{0.995}\frac{s}{\sqrt{n}})$$

$\Leftrightarrow$

$$\frac{\bar{q} - Q_n}{s/\sqrt{n}} \geq -t_{0.995} \qquad (1)$$

Let $T = \dfrac{\bar{q} - Q_n}{s/\sqrt{n}}$, so here the equation $\bar{q} \geq (Q_n - \lambda s)$ is equivalent to a test based on the statistic T.

Let Q be a random variable and denote the package contents of the products. Q follows a normal distribution. Let $\bar{Q}$ denote the unknown population mean and n denote the sample size. Then $Q_1, Q_2, ..., Q_k$ is a random sample chosen from the products and $\bar{q}$ is the sample mean. From Theorem 2.3.2 we know $\bar{q}$ follows a normal distribution. Note that s is the sample standard deviation and Q_n is the null value of the mean package contents (the net quantity). We can conclude that T follows a Student-T Distribution and the test is basically a T-test.

2.1 Confidence Interval for $\bar{Q}$

At the very beginning, we note that the notation $t_{0.995}$ is the same as the value $t_{0.005}$ defined on the slides. By some calculation using the data in Table 4 of the material,

$$t_{0.995,9} = 1.028 \times \sqrt{10} = 3.251,$$

while $t_{0.005,9} = 3.250$ by the definition on the slides, so they are equal and have the same meaning. It is similar for other data. In the section below, we use the notation $t_{0.995}$ in the material but it actually represents $t_{0.005}$ on the slides.

Recall the definition of $t_{0.995}$, we know that the probability that T is greater than or equal to $-t_{0.995}$ is 99.5%. In other words, the probability that equation (1) is true is 99.5%.

On the other hand, we can rewrite equation (1) as

$$Q_n \;\leq\; \bar{q} + t_{0.995}\frac{s}{\sqrt{n}}$$

so we can say the probability that Q_n is less than or equal to the right-hand side of the equation is 99.5%. But the right-hand side of the equation is exactly the 99.5% upper confidence bound (or the upper one-sided confidence interval) on $\bar{Q}$. Hence, the original equation $(*)$ is true if and only if Q_n lies in the 99.5% upper confidence interval of the mean package content $\bar{Q}$.

2.2 Hypothesis Test

To see the logic of the hypothesis test hidden behind equation $(*)$, the key is to find the null hypothesis H_0. We have two ways to find it.

The first way is to connect the confidence interval of the null value Q_n and the sample mean $\bar{q}$. We know that the sample mean $\bar{q}$ lies in the critical region for α if and only if the null value Q_n does not lie in a $100(1-\alpha)\%$ confidence interval for the population mean Q. That implies that Q_n lies in the 99.5% upper confidence interval of $\bar{Q}$ if and only if $\bar{q}$ does not lie in the critical region for 0.5%. That means we accept the null hypothesis, so the null hypothesis is $H_0 : \bar{Q} \geq Q_n$.

The second way is to compare the case with a general T-test. For a general T-test, $T_{n-1} = \dfrac{\bar{X} - \mu_0}{s/\sqrt{n}}$, we reject $H_0 : \mu \geq \mu_0$ at significance level α if $T_{n-1} < -t_{\alpha,n-1}$. For equation

(1), however, T is greater than or equal to $-t_{0.995}$ (the reason why the subscript $n-1$ is missing is that the people making the given table have taken it into consideration). That implies we accept H_0 where $H_0 : \bar{Q} \geq Q_n$.

Based on the H_0 we confirmed, the logic of the hypothesis test is as follows.

We want to test whether the products have package contents corresponding to the net quantity marked on the packaging, so we set the hypothesis $H_0 : \bar{Q} \geq Q_n$. We want this test to have 99.5% of significance, namely if H_0 is true, there is only 0.5% of probability that it is rejected. Then we randomly selected the sample based on the quantity of these products and calculate the sample mean. We get the values of $t_{0.995}/\sqrt{n}$ from the table and try to verify equation $(*)$. If it is false, equation (1) is false so we reject H_0 and reject the products. If it is true, that means equation (1) is true so we accept H_0 and accept the products.

3 Single Attribute Sampling Plan for Acceptance Sampling

3.1 The sampling scheme

In 5.3.1 there are three rules to decide whether a bunch of products should be rejected. The sampling document was not the first one to have brought forward these rules. Back in 1981, the British government has established the three packer's rule[2], which generally coincides with those in 5.3.1. The idea and statistics behind table 4 we wish to discuss in this section was not a new idea either. In the standard document ANSI/ASQ Z1.4-2003, published by American Society for Quality[8], a table(Fig. 1) was offered to provide guidance in terms of single sampling for different acceptance quality limit. The numbers we need to discuss is in the column where the Acceptance Quality Limit is 2.5%.

3.2 The significance

The sampling plan in 5.3.2.2 of JJF1070-2005 essentially says that the probability for rejecting a qualified inspection lot containing no more than 2.5% bad quality products does not exceed 5 %. To put in a simpler manner, the probability of sending stuff back even though the lot meets the standard is no more than 5%. By definition, this is the consumer's risk α, also can be interpreted as the Type I error. We can thereby conclude the significance for Type I error is 5%.

Referring to the rule in 5.3.3 which says that the test can be 90 % accurate in detecting the deficiency in average value and the lot containing 9% of items of bad quality. The significance of the Type II error is evaluated by the probability of accepting a lot even though it is bad. This can be calculated by $1 - 90\% = 0.1$.

Nevertheless, we still need further calculation that these rules actually conforms to the statistic in Table 4, and a detailed discussion will be presented in the upcoming section.

Table II-A—Single sampling plans for normal inspection (Master table)

(See 9.4 and 9.5)

Sample size code letter	Sample size	Acceptance Quality Limits, *AQLs*, in Percent Nonconforming Items and Nonconformities per 100 Items (Normal Inspection)																									
		0.010	0.015	0.025	0.040	0.065	0.10	0.15	0.25	0.40	0.65	1.0	1.5	2.5	4.0	6.5	10	15	25	40	65	100	150	250	400	650	1000
		Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re
A	2														↓	0 1		↓	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	30 31
B	3													↓	0 1	↑	↓	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	30 31	44 45
C	5												↓	0 1	↑	↓	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	30 31	44 45	↑
D	8											↓	0 1	↑	↓	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	30 31	44 45	↑	
E	13										↓	0 1	↑	↓	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	30 31	44 45	↑		
F	20									↓	0 1	↑	↓	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	↑	↑	↑			
G	32								↓	0 1	↑	↓	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	↑						
H	50							↓	0 1	↑	↓	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	↑							
J	80						↓	0 1	↑	↓	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	↑								
K	125					↓	0 1	↑	↓	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	↑									
L	200				↓	0 1	↑	↓	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	↑										
M	315			↓	0 1	↑	↓	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	↑											
N	500		↓	0 1	↑	↓	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	↑												
P	800	↓	0 1	↑	↓	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	↑													
Q	1250	0 1	↑	↓	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	↑														
R	2000	↑		1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	↑															

↓ = Use the first sampling plan below the arrow. If sample size equals, or exceeds, lot size, carry out 100 percent inspection.
↑ = Use the first sampling plan above the arrow.
Ac = Acceptance number.
Re = Rejection number.

Figure 1: Single Sampling Plan For Normal Inspection, American Society For Quality[8]

3.3 Two approximations to single sampling plan

To start with we define the three-tuple for Hypergeometric p distribution(N,n,c), where d lies between 1 and c. The probability is then calculated by

$$P(c, N, n, r) = \sum_{d=0}^{c} \frac{\binom{r}{d} \cdot \binom{N-r}{n-d}}{\binom{N}{n}}$$

N is the total number of the population(products); lot size
c is the number of defectives in the sample
n is the sample size
r is the number of defectives/bad quality items in the population
P is the probability of passing the test.

Two possible approaches can be used to approximate this distribution when N is large(Or to be specific, $\frac{n}{N} \leq 0.1$). Here we have introduced another parameter p which is an estimate for $\frac{r}{N}$. The first approach is based on Binomial Distribution

$$P = \sum_{d=0}^{c} \binom{n}{d} \cdot p^d \cdot (1-p)^{n-d} \tag{1}$$

The second approach is based on Poisson

$$P = \sum_{d=0}^{c} \frac{np^d \cdot e^{-np}}{d!} \tag{2}$$

However, if we want to approximate hypergeometric to Poisson or binomial, a crucial condition to satisfy is $\frac{n}{N} \leq 0.1$. This following discussion will revolve around this constraint.

To take a step further, we have to first treat two concepts properly. Namely, the producer's risk(α) and the consumer's risk(β). The producer's risk essentially measures the probability that the test rejects the products while the product is acceptable, which damages the benefit of the producer. The consumer's risk is the probability failing to reject the unqualified products which in this way the rights for the consumer is not guaranteed. It is important to note that α and β are two conflicting and mutually exclusive concepts. In other words, by reducing the consumer's risk will almost surely increase the producer's risk, and vice versa. So in practice, it is important to find an appropriate pair of these values so that both the consumer's rights and producer's cost is taken into proper consideration. Thereby, an appropriate sampling plan should take good care of the producer's risk and consumer's risk simultaneously.

3.4 Discussion for the number "0,0,1,3,5,7"

As it is introduced, both Poisson distribution and Binomial distribution can be used to approximate the Hypergeometric distribution. However, a crucial premise is that the value for n/N should be less than 0.1. Referring back to the Form 4 that we need to further investigate, we discuss how "3", "5", "7" are derived since some of the N's and n's in the corresponding rows suffice the condition $n/N \leq 0.1$

3.4.1 The derivation of "3","5","7" using Binomial Distribution

To recap, we approximate the hypergeometric distribution by the following binomial distribution

$$P = \sum_{d=0}^{c} \binom{n}{d} \cdot p^d \cdot (1-p)^{n-d}$$

We note that the parameter p here is an approximation which stands for the fraction of bad quality items in the lot(whole population). Given that a lot is good, if p is less than or equal to AQL(Acceptable Quality Level), which can be defined as if there are p percent defective in the lot, there is 1-α probability of acceptance. To simplify, the probability of accepting a qualified lot without exceeding AQL is 1-α.

The counterpart for AQL(Acceptable Quality Level) is the LTPD(Lot Tolerance Proportion Defective) or the worst level that a consumer can tolerate[5]. The consumer wishes to detect with 1-β certainty that a lot contains p percent of defectives, where p equals to the LTPD. In other words, the probability of LTPD quality is the consumer's risk(β).[5]. An intuitive explanation toward this concept(LTPD) is that it resembles the fraction of bad quality that the consumer would like to reject with $(1-\beta)*100$ percent certainty. We denote the value for AQL and LTPD as p_0 and p_1 respectively. Then we can associate them with α and β. Namely,

$$\begin{cases} 1-\alpha = \sum_{d=0}^{c} \binom{n}{d} p_0^d \cdot (1-p_0)^{n-d} (*) \\ \beta = \sum_{d=0}^{c} \binom{n}{d} p_1^d \cdot (1-p_1)^{n-d} (**) \end{cases}$$

We denote f(c,p,n)=$\sum_{d=0}^{c} \binom{n}{d} p_0^d \cdot (1-p_0)^{n-d}$. A little transformation soon yields,

$$\begin{cases} \alpha = 1 - f(c, p_0, n)(*) \\ \beta = f(c, p_1, n)(**) \end{cases}$$

In the case of JJF1070-2005, p_0=2.5%(5.3.2.2). The number 2.5% conforms to the rules in 5.3.2.2 which says the probability of rejecting a lot containing no more than 2.5% defectives is no more than 5% and the definition for AQL which is defined as the p (percent defective)with a 95% probability of acceptance. In a similar manner, we can conclude p_1=9%(5.3.3).

We notice that α, β, p_0, p_1 are already determined in this case. Therefore, the solution tuple (n,c) can then be uniquely determined. The domains for n and c means that they can not take on arbitrary values, but can only take on positive integer values. This makes the approach of obtaining a numerical solution approach infeasible. Despite the infeasibility, we still want an approximation which can nonetheless provide guidance for the process of single attribute sampling[3].

To start with, we examine the scenario when the sample size is equal to 125. Using Mathematica, we have the following result(Fig. 2)

```
f[c_, p_, n_] := Sum[Binomial[n, d] * p^d * (1 - p) ^ (n - d),
                 |…    |二项式系数
  {d, 0, c}]

In[164]:= DiscretePlot[{1 - f[c, 0.025, 125], f[c, 0.09, 125]},
          |离散图
  {c, 5, 8}, PlotLabel → SampleSize : 125,
             |绘图标签
  GridLines → Automatic,
  |网格线      |自动
  PlotMarkers → Automatic,
  |绘制点的标记  |自动
  LabelStyle → Directive[Bold, Orange], Joined → True]
  |标签样式      |指令      |粗体  |橙色     |连接点   |真
```

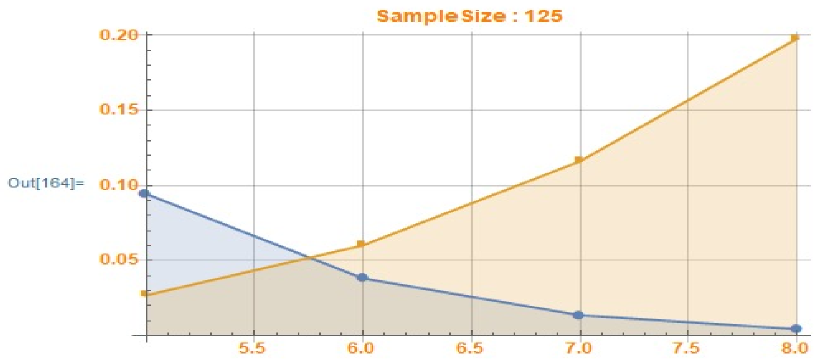


Figure 2: Sample size = 125, plotted in Mathematica

It may be to some extent confusing and misleading to connect lines between those discrete points. However, for the sake of clarity, it is still worthwhile doing so. As can be inferred from the graph and the code, the orange line is describing the value for f[c,0.09,125], our goal is to find a value that is nearest to 0.1. In other words, to find a solution so that (**) holds, we wish to find a c that is as close to 0.1 as possible. Similarly, the blue curve is just the right-hand side (*), and we would like to find a value that is as close to α, which is 0.05. Looking at the graph, when c=6, the blue point is closer to 0.05 than that when c=7, while when c=7, the orange point is closer to 0.05 than that when c=6. Taking a step back, and instead of looking at the plot graphically, we look at the idea of consumer's risk, and producer's risk properly.

We know that α and β are two variables negatively correlated. It is almost surely impossible to design a sampling plan that can benefit both the producer and the consumer. Therefore, certain compromise has to be made either from the consumer or the producer.

In this case, the consumer has to make a compromise because when c=6, f(6,0.09,125)≤0.1. The producer would argue that consumer should undertake more risks by increasing the acceptance number c so that it is closer to the default setting. It can be inferred that this sampling plan is in favor of the benefits of the producer rather than the consumer. Nevertheless, the sampling rule will only suffice the benefit of the producer to a certain extent so that the benefits are not damaged too much as well. We thereby postulate the following sampling rule which is applied to determine c.

Sampling Rule: For fixed n, the acceptance number c is found by the smallest integer such that

$$\begin{cases} \alpha \geq 1 - f(c, p_0, n) \\ \beta \leq f(c, p_1, n) \end{cases}$$

In the case when sample size is equal to 80, and 50. We have the following two plots(Fig. 3&4)

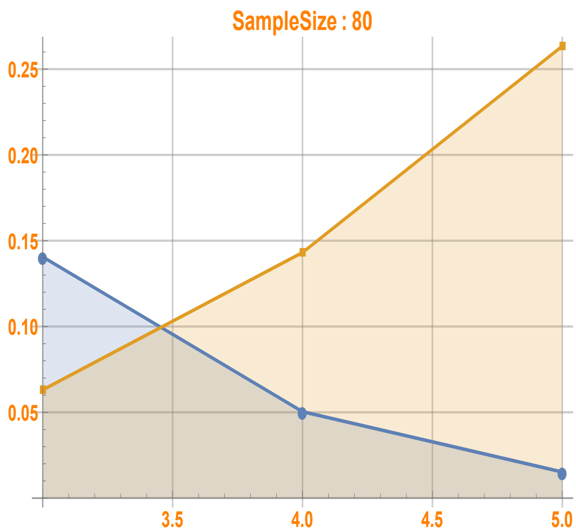


Figure 3: Sample Size 80, Mathematica

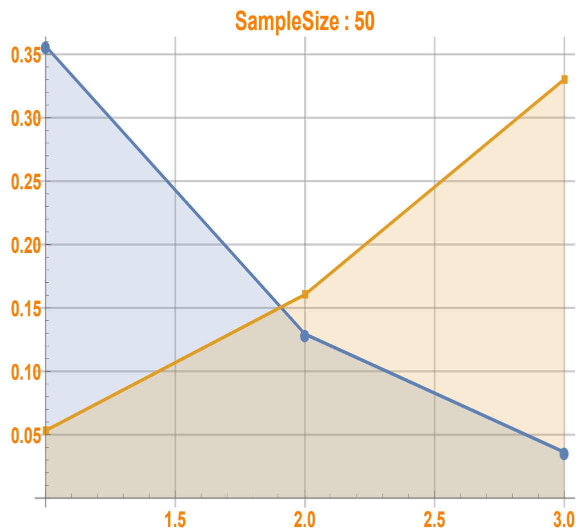


Figure 4: Sample Size 50, Mathematica

With similar codes in Mathematica, we obtain the two plots corresponding to the sample size of 80, and 50. In the left graph, the smallest c satisfying the sampling rule is c=5(when c=4, 0.05≤1-f(c,0.025,80)), and the smallest c corresponding to the scenario on the right hand side is c=3(when c=2, 0.05≤1-f(c,0.025,50)). At this point, we would like to make the comment that binomial distribution is not the only distribution that obtains this result. A similar approximation can also be obtained in Poisson.

3.4.2 The derivation of "0,0,1"

3.4.2.1 Why or why not Approximate

In general, we have to separate our discussion for single attribute sampling plan into two circumstances: the case when the lot size is relatively large where we can use the binomial or Poisson approximation, and when the lot size is finite where we should use hypergeometric distribution. The operating characteristic curves that we used a binomial distribution to construct is often referred to as the Type B OC curve, while the curves we use a hypergeometric distribution to construct is called Type A OC curve.[6] Type-A OC curves are concerned with the finite size of the lot and can be used to examine consumer's risk since a consumer wishes to know concretely(in a finite size) what is the risk of accepting a bad lot.[6] Type-B OC curves can be used to examine a producer's risk since a producer generally wants to know the average quality that his or her stream of lots has to reach so that it would not be rejected.

In practice, we would like to approximate it because it is easier for calculation. Using a binomial distribution or a Poisson distribution to approximate, we simply neglect the variable N, which greatly simplifies the calculation. In fact, we have to admit we are being quite "sloppy" and actually did some "cheating" in the previous discussion because for example in the fourth column, when N takes the value 100, and n=50, $\frac{n}{N} \leq 0.1$ does not hold. Nevertheless, we still went ahead and did the calculation.

3.4.2.2 The argument for the first zero

The argument is rather straightforward and intuitive. Since the sample size is equal to the lot size, a simple contraposition argument is that if c is not equal to zero, then the consumer's risk is bigger than 0.1 which obviously violates the sampling rule.

3.4.2.3 The Dilemma for using hypergeometric distribution

The dilemma encountered in dealing with the data in the second and the third column is that on one hand simply closing our eyes neglecting the condition to use a binomial approximation violates the rule, while on the other hand dealing with hypergeometric distribution is kind of tricky. To treat a hypergeometric distribution function properly, we have to consider three different variables, N,n,c. To analyze the relation of the tuple (N,n,c) quantitatively given p_0, and p_1, we have to examine the graph for P(c,N,n,p*N) given p, where c, N, n takes on discrete integer values. To completely present how P works, rigorous analysis is desired. However, the mathematical analysis needed to treat P properly is beyond the scope of this project, and results should be presented in a 3-dimensional surface, which is impossible to implement at the moment. Plus, it is actually unnecessary because the sampling plans are, in fact, a process of negotiation between the producer and the consumer. Therefore, instead of presenting mathematical formulas, graphs and possible reasoning why the values c actually works will be presented.

3.4.2.4 The "0" and "1"

As it is mentioned in the previous section, the goal of this section is not to interpret is not to give a concrete and rigorous mathematical analysis nor to give a recipe of finding a sampling plan that can suit different Ns. Rather, we wish to verify that c=0, and c=1 is the optimal choice in this case. In the following discussion, we assume that the lot size and the sample size is already fixed and we wish to find an appropriate value for c. Using the following

Mathematica command

$$f[c,N,n,d] := Sum[Binomial[d,i]*Binomial[N-d,n-i]/Binomial[N,n],\{i,0,c\}]$$

W then plot two plots in mathematica(Fig. 5&6). p On the left is the mathematical plot

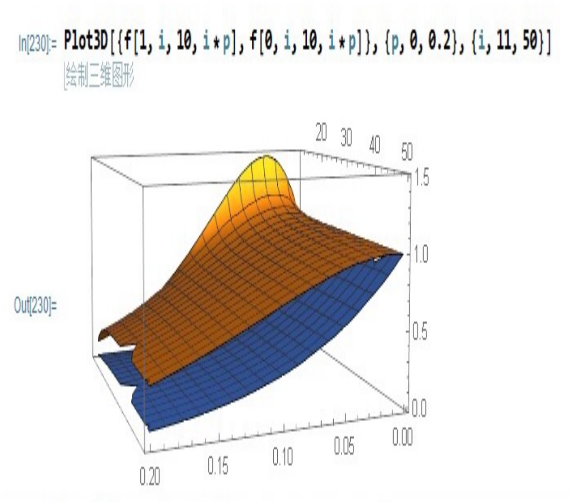


Figure 5: Sample Size 10, Mathematica

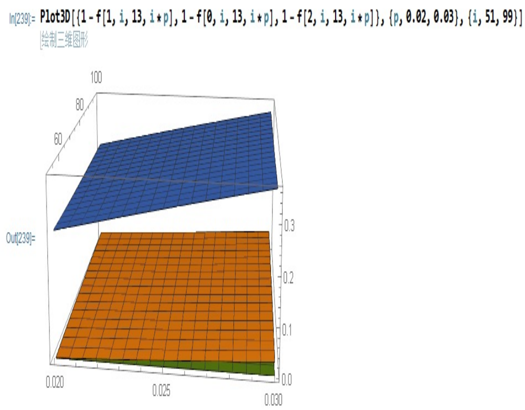


Figure 6: Sample Size 13, Mathematica

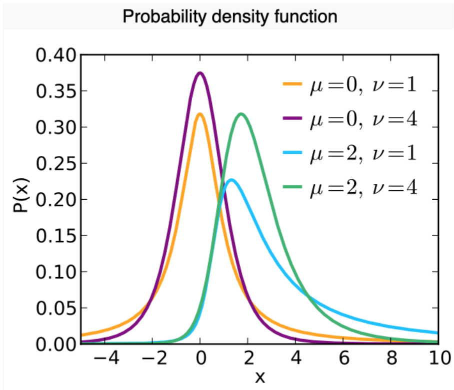
where we plot the hypergeometric distribution for the case where the sample size is 10. Instead of plotting a OC curve, we plot an OC "surface". Pointing to the left direction is the value for p, which equals the value $\frac{r}{N}$. The one lying below is the scenario when c=0, and we examine the line where p=0.09(In this case we are generally interested in the consumer's risk or the LTPD value). In general, we wish that the value is somewhat less than 0.1, or at least close to 0.1. Therefore, we can then argue that c=0 is an more optimal option because you can get a lower consumer's risk, which is what we desired in this case. Lastly, we would like to comment on the effect of the little heave on the top. The i*p does not always yield integer value, so there some sort of rounding in this case and cased the deviation.

On the right is the circumstance when c=0, c=1, and c=2. In this case, considering the consumer's risk does not yield a proper result,because both c=0, and c=1 and fairly greater than 0.1. So we consider the producer's risk in this case, and we hope to get risk as close to 0.05. at the point where p=0.025, and we see that c=1, which is the curve fits the best. Although c=2 would also fit(which is the surface under it), but we think that in this case it is not proper because the risk for the consumer is too large, and we want to find the smallest c so that the producer's risk is minimized. It can be read off from the graph that c=1 is the optimal value(closest to 0.05 and smallest).

4 Non-central T distribution

4.1 Derivation of the OC Curve

A non-central T distribution is a generalization of student T-distribution with a noncentrality parameter. It describes how T is distributed when the null is false.The figure of probability density function of the noncentral T distribution is shown below(Fig. 7).



p

Figure 7: The PDF for non-central T distribution, Wikipedia[9]

For the analysis of package contents, we need to use the non-central T distribution when doing the T test mainly because of the fact that this is a one-sided T test and in our hypothesis $\mu \neq \mu_0$ anyway. That is to say, $\overline{q} \neq Q_n$. Thus if we want to derive a formula for the OC curve of the T test, we need to use the non-central T distribution.

Our discussion is based on *Applications of the Noncentral tCDistribution* by F.W. Scholz[10].

Denote $T_{f,\delta} = \frac{Z+\delta}{\sqrt{V/f}}$ where f is the degree of freedom, δ is the non-central parameter, Z denote the standard normal variables and V are chi-squared variables.

Denote the cumulative distribution function of $T_{f,\delta}$ as $G_{f,\delta}(t) = P(T_{f,\delta} \leq t)$ For the T-test, we reject at significance α $H_0 : \overline{q} - as \geq Q_n$ $H_1 : \overline{q} - as < Q_n, \ a = -\lambda$
We have that β is the probability of making type II error. In the case of this problem, we want the hypothesis H_0 to be true while reject H_1. That is to say, here β should be the probability of rejecting H_0

$$\beta = P[\frac{\overline{q} - Q_n}{S/\sqrt{n}} > t_{\alpha,f}]$$

$$\frac{\overline{q} - Q_n}{S/\sqrt{n}} = \frac{\sqrt{n}(\overline{q} - Q_n)}{s} = \frac{\sqrt{n}(\overline{q} - \mu_0)/\sigma - \sqrt{n}(Q_n - \mu_0)/\sigma}{s/\sigma} = \frac{Z+\delta}{\sqrt{V/f}}$$

Where $f = n-1, \delta = -\sqrt{n}(Q_n - \mu_0)/\sigma, \ V = \frac{s^2 f}{\sigma^2}, \ s/\sigma = \sqrt{V/f}$
Thus

$$\beta = P[\frac{\overline{q} - Q_n}{S/\sqrt{n}} > t_{\alpha,f}] = P[\frac{Z+\delta}{\sqrt{V/f}} > t_{\alpha,f}] \leq 1 - P[\frac{Z+\delta}{\sqrt{V/f}} = t_{\alpha,f}] = 1 - G_{f,\delta}(t_{\alpha,f})$$

$$1 - \beta = G_{f,\delta}(t_{\alpha,f})$$

$V = \frac{s^2 f}{\sigma^2}, \ f = n-1, \ \delta = -\sqrt{n}(Q_n - \mu_0)/\sigma = \sqrt{n}d \ , \ \ d = \frac{\mu_0 - Qn}{\sigma}$
Thus the formula for the OC curve is $G_{f,\delta}(-t_{\alpha,f})$ where $f = n-1, \delta = \sqrt{n}d, \ d = \frac{\mu_0 - Q_n}{\sigma}, \alpha = 0.995$
The notation on the project supporting material is a bit different from the notation on Mathematica and slides. Here we use the notation on the project support materials, which means $\alpha = 0.995$ just means the $\alpha = 0.005$ in Mathematica.

5 OC curve for the T distribution

Mathematica has built in function for Non Central T distribution. Using mathematica we plot the OC curve for n=10, 13, 50, 80, 125(Fig. 8).

```
In[29]:= Plot[
         绘图
  Table[1 - CDF[NoncentralStudentTDistribution[n - 1, Sqrt[n] * d],
  表格      ···  非中心学生t分布                      平方根
    InverseCDF[StudentTDistribution[n - 1], 0.005]], {n, {10, 13, 50, 80, 125}}],
    逆累积分布函数 学生t分布
  {d, 0, -2}, Filling → Axis,
              填补       轴
  PlotLabels → Placed[{"10", "13", "50", "80", "125"}, {Scaled[1], Before}],
  数据绘制标签   放置                                      比例坐标     在之前
  AxesLabel → {"d", "beta"}]
  坐标轴标签
```

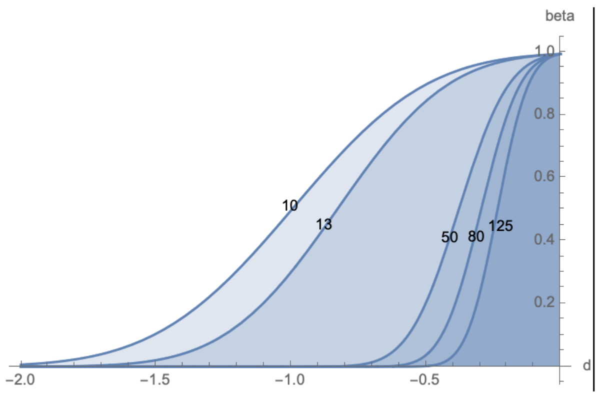


Figure 8: OC curve for the T-Test with different sample sizes, Mathematica

Some comment about power and OC curve:[3].

1. For the same d, we see that when n increases, β becomes smaller, that is to say, its power becomes larger. In other words, we can say that the probability of accepting H_0 converges to zero in a pointwise manner. This means that when the sample size n increases, its ability to discriminate between good and bad lots increases. But in practice, one should address the practical significance issue of any difference μ-μ_0 and weigh that against the cost of a large sample size.

2. For the same n, we see that when d is much closer to 0, power decreases, this means that it is more difficult to distinguish the good and bad lots.

3. Decreasing the acceptance number actually shifts the OC curve to the left, but also increases the protection of the plan. That is, for the same n, d, if the probability of rejecting H_0 decreases, then power increases. That is, the probability of making the correct decision increases, it becomes easier to make the correct decision.

4. There will not have a significant effect on the OC curve unless the sample size is large ($N/n < 10$)

Thus we can see that the OC curve actually allows us to see the level of protection provided by a sampling plan [3].

When there is at least one standard deviation in package contents, this is means $d = -1$, we estimate the probability directly from the OC curve(Fig. 9).

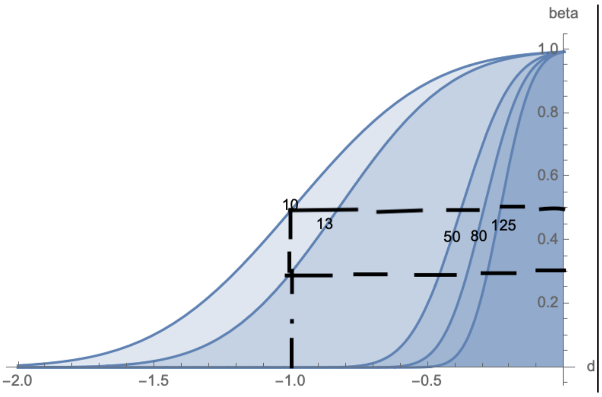


Figure 9: OC curve where d=-1 examined, Mathematica

Thus we see that for n = 50,80, 125, the probability of accepting H_0 is almost zero. That is to say, the probability of detecting a shortfall is almost 1. For n = 13, the probability of detecting a shortfall is about $1 - 0.3 = 0.7$. For n = 10, the probability of detecting a shortfall is about $1 - 0.5 = 0.5$. The calculation is based on Mathematica(Fig. 10).

```
In[3]:= Table[1 - CDF[NoncentralStudentTDistribution[n - 1, Sqrt[n]],
        表格      |…  非中心学生t分布                          平方根
          -InverseCDF[StudentTDistribution[n - 1], 0.005]], {n, {10, 13, 50, 80, 125}}]
           逆累积分布函数 学生t分布

Out[3]= {0.504107, 0.700644, 0.999989, 1., 1.}
```

Figure 10: Mathematica code to calculate the shortfall, Mathematica

6 Summary and Discussion of Section 5.3.2

6.1 Summary

The section 5.3.2 provides the significance level of T1 risk test, namely the producer's risk test. There are two kinds of significance level for the T1 risk, respectively introduced in 5.3.2.1 and 5.3.2.2.

- For 5.3.2.1, the test is for mean package content. If the average actual quantity of the product lot equals the nominal quantity, the probability for the lot to be rejected is no more than 0.5%. The method is one-sided T-test.
- For 5.3.2.2, the test is for the net quantity of a unit product. The practical meaning is that for the inspection lot containing no more than 2.5% inadequate products, the probability of the product lot being rejected should not exceed 5%.

6.2 Discussion

6.2.1 Comprehension

To help clarify, we will first restate the whole control criterion according to Table 4 in JJF1070-2005.

1. The average actual quantity should be no less than the difference of nominal net quantity and the modification value of the average actual quantity of the products lot.
2. Then number of T1-shortage product in the sample should be no more than 2.5% of the sample size.
3. There is no T2-shortage product in the sample.

The three regulations can be divided into two categories. The first event (1) concerns the average actual quantity of the inspection lot (denoted by A). Another two events (2, 3) focus on the net quantity of a unit product in pre-packages with fixed content (denoted by B). Then the two subsections in 5.3.2 correspond to the significance level of event A and event B respectively. And we believe this is a reasonable classification. It clearly distinguish between the producer's risk (5.3.2) and consumer's risk(5.3.3). It consider both the mean package contents and content for unit product in the standard.

6.2.2 Verification

Then we will give verification that the given standards in 5.3.2 can result in the criterion in Table 4 in the JJF1070-2005.

6.2.2.1 Average Actual Quantity

As we have verified in question one, λ s is related to a confidence interval for the mean package contents. We will derive how is the standard 5.3.2.1 related to this result.

Since the actual quantity is a sample with a large sample size, we can say that $\dfrac{\overline{q}-\mu}{s/\sqrt{n}}$ follows a T_{n-1}-distribution. To ensure that when the average actual quantity of the product lot equals to nominal quantity, the probability for the lot to be rejected is no more than 0.5%, we use a T-test to realize the request. We let $H_0 : \mu \geq Q_n$. It will be rejected at a significance level $\alpha_n = 0.5\%$ if $T_{n-1} > t_{\alpha_p,n-1}$, where $T_{n-1} = \dfrac{\overline{q}-Q_n}{s/\sqrt{n}}$. Therefore,

$$\frac{\overline{q}-Q_n}{s/\sqrt{n}} \geq t_{0.005}$$

$$\bar{q} \geq Q_n + \frac{t_{0.005}}{\sqrt{n}} s$$

$$\bar{q} \geq Q_n - \frac{t_{0.995}}{\sqrt{n}} s$$

Then we let $\lambda = \dfrac{t_{0.995}}{\sqrt{n}}$, we got

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

which correspond to the request in Table 4. So the significance level α works well.

6.2.2.2 Net Quantity of a Unit Product

We apply the accepted number of Type I inadequate products (0,0,1,3,5,7) respectively to each conditions of sample size. We will check whether the calculated probability is no more than the given significance level α_μ. The formula used is derived from hypogeometric distribution as shown in question 2. The calculation is completed using the following Mathematica command

$$f[c, N, n, d] := Sum[Binomial[d, i] * Binomial[N - d, n - i]/Binomial[N, n], \{i, 0, c\}],$$

where c is the accepted number of Type I inadequate products, N is the inspection lot size and n is the sample size. $f[c, N, n, d]$ is the probability we need. Denote $f[c, N, n, d]$ by p. And we aim to obtain $p \geq 1 - \alpha_\mu = 0.95$. All the results are shown in the following table.

N	Lower Bound	Upper Bound
1-10	0.975	0.975
11-50	0.561	0.756
51-99	0.988	0.971
100-500	1.000	0.972
501-3200	0.993	0.986
> 3200	0.988	/

Table 1: The calculated significance level.

We can find that all the calculated significance levels are greater than 0.95 except for N between 11 to 50.

6.2.3 Criticism and Suggestion

From the calculation above, we can conclude that there exist some problems since the significance level artificial standards. Sometimes they could not satisfy all the conditions in real industrial manufacture. For example, when the inspection lot size is between 11 to 50, the producer's risk is between 24.4% to 43.9%, which is obviously too large and not reasonable.

Our suggestion is to set a special criteria for the product size between 11 to 50, thus eliminating the producer's risk. We believe the condition worth another test because the difference of the present probability and the given 5% is large enough to cause a great loss for the producer. We are also looking forward to a better solution from the relevant department through adjusting the significance level settled after rigorous calculation and thinking.

7 Prepackaged Food Sample Test

In the test, we obtained a total number of 16 chocolate wafers with a nominal quantity Q_n of 17g. According to the requirements described in Table 4 in [11], the number of samples should be 10 for a total number of products between 11 to 50. So we randomly select 10 of them to produce the test. From Table 3 in [11], the tolerable Inadequate T is

$$T = 9\% \times Q_n = 1.53g \approx 1.5g. \tag{3}$$

We measured their total weight GW and the tare weight P. The average tare weight is

$$\overline{P} = \frac{1}{10}\sum_{i=1}^{10} P_i = 0.42g. \tag{4}$$

We can see that

$$\overline{P} = 0.42g \leqslant 10\% \times Q_n = 1.7g \tag{5}$$

So we use $\overline{P}$ as the tare weight for all samples. Then we can calculate the Net quantity q and the deviation D for all these samples.

$$q = GW - \overline{P} \tag{6}$$

$$D = q - Q_n \tag{7}$$

The data of the 10 samples are listed in Table 2.

Table 2: Data of chocolate wafers net quantity test

Nominal quantity Q_n [g]	17	Tolerable Inadequate T [g]	1.5
Average tare weight $\overline{P}$ [g]	0.42	Tare sample quantity n_P	10
Total weight GW [g]	Tare weight P [g]	Net quantity q [g]	Deviation D [g]
17.45	0.42	17.03	0.03
17.97	0.41	17.55	0.55
17.89	0.42	17.47	0.47
18.76	0.42	18.34	1.34
17.84	0.42	17.42	0.42
19.09	0.42	18.67	1.67
20.14	0.42	19.72	2.72
20.00	0.43	19.58	2.58
18.85	0.42	18.43	1.43
19.28	0.44	18.86	1.86

We can see that all the deviation D is positive, which means that the net qualities of all the samples are larger than the nominal quantity. So the number of samples with type T1 shortage is 0. These chocolate wafers pass the net quantity test by our sample test.

Then we plot this non-central T distribution of the test. The freedom is calculated as

$$f = n - 1 = 9. \tag{8}$$

The noncentrality parameter is calculated as

$$\delta = \sqrt{n}\frac{\mu_0 - Q_n}{\sigma} \tag{9}$$

The sample standard deviation is calculated as

$$s = \sqrt{\frac{1}{10-1}\sum_{i=1}^{10}(q_i - \overline{q})^2} = 0.929g \tag{10}$$

From [12], The variance for Student T distribution with freedom 9 is

$$\sigma^2 = \frac{f}{-2+f} = \frac{9}{7} \tag{11}$$

$$\sigma = \sqrt{\frac{9}{7}} = 1.134 \tag{12}$$

$\lambda = 1.028$ for n=10. So the noncentrality parameter is

$$\delta = \sqrt{n}\frac{\mu_0 - Q_n}{\sigma} = 1.484 \tag{13}$$

The T-distribution curve of the chocolate wafers is plotted as

```
In[3]:= Plot[PDF[NoncentralStudentTDistribution[9, 1.484], x] // Evaluate, {x, -5, 15},
        绘图    ⋯  非中心学生t分布                                              计算
         PlotRange → {0, 0.4}, Filling → Axis]
         绘制范围                  填补       轴
```

Out[3]=

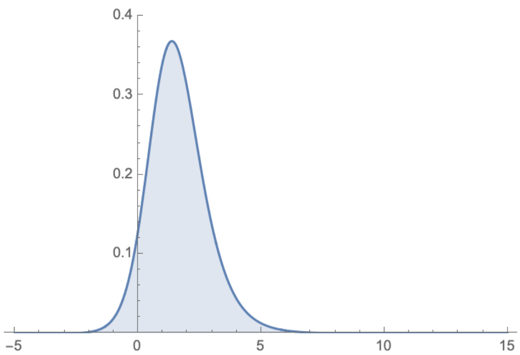


Figure 11: T-distribution of the chocolate wafers, Mathematica

From this figure we can see that most chocolate wafer should be 1 to 3 grams heavier than the nominal quantity, which is in accordance with our experimental data. However, there is no data points showing that there exists chocolate wafers 3 grams heavier than the nominal quantity. But in the non-central T distribution, these chocolate wafers should take a certain portion. This may be a result of our small sample size.

8 Final comments and insights

Besides the methods introduced in this project, there are also other statistics tools such as the Larson Nomogram[7] which can also be used to read off producer's risk/consumer's risk from AQL/LTPQ. In this project, we were unable to give a recipe for finding the sample size for a given lot size. However, we speculate that the process of choosing an appropriate sample size involves balancing both the consumer's risk and the producer's risk in a subtle manner. We would also like to note that non-central distribution T test lay an important basis for variable attribute sampling whose discussion is beyond the scope of this project.

9 Appendix

[1]陶亚明,陈黎明.定量包装商品净含量计量检验及管理对策[J].江苏现代计量,2015,(1):32-36.

[2]The stationery office of Dublin. Guidance manual for packers and importers. The minister for industry, Commerce and Tourism in conjunction with the Packaged goods Regulations, 1981,(39):7-8

[3]Raul Soto. Sampling Plans that Result in Statistical Confidence. http://www.cbinet.com /sites/default/files/files/Soto_Raul_UPDATED %20Session%2013_Raul_Sampling%20Plans% 20.pdf

[4] D. F. Bray , D. A. Lyon, and I. W. Burr (1973) Three Class Attributes Plans in Acceptance Sampling, Technometrics, 15:3, 575-585

[5] http://wps.prenhall.com/wps/media/objects/7117/7288732/65767_28_SuppG.pdf

[6]Schalk W. Human, Marien A. Graham(2014). Average Run Lengths and Operating Characteristic Curves,Wiley StatsRef: Statistics Reference Online, 1-16p.

[7]http://web.tecnico.ulisboa.pt/mcasquilho/acad/cq/qfund/Q_AS_UTexD.pdf

[8]American Society For Quality. American National Standards ANSI/ASQ Z1.4-2003. 600 North Plankton Avenue Milwaukee, Wisconsin 53203

[9]Wikipedia. Noncentral t-distribution. Wikipedia, the free encyclopedia, 2014. https://en.wikipedia. org/wiki/Noncentral_t-distribution [Online; accessed 05-April-2019].

[10]F. W. Scholz. Applications of the Noncentral t-Distribution. Boeing Computer Services. http://faculty.washington.edu/fscholz/Reports/nctnewrpt.pdf [Online; accessed 05-April-2019].

[11] 全国法制计量管理计量技术委员会/定量包装商品净含量工作组. 定量包装商品净含量计量检测规则(Rules of metrological testing for net quantities of products in prepackages with fixed content). Technical Re- port JJF 1070-2005, 中国计量出版社, 2005.
http://www.ahhz.gov.cn/CountyOpennessContent/show/ 1078758.html [Online; accessed 22-March-2019].

[12]Wolfram, Wolfram Mathematica Online, http://mathematica.wolframcloud.com/app/ [Accessed at 6 April 2019].

[13] 邵振威, 尚贤平, 杨眉. 关于JJF1070-2005《定量包装商品净含量计量检验规则》中抽样原理的探讨(Discussion on sampling principle in JF 1070-2005 "Quantitative net content of packaged goods rules of Metrological Inspection"). China Academic Journal Electronic Publishing House, 浙江计量科学研究院, 2012. http://kns.cnki.net/KCMS.

[14] Avery Weigh-Tronix. Average Weight Legislation: A quick Reference Guide.

https://www.averyweigh-tronix.com/en-GB/News/Weighing-tips/General-weighing-tips-and-guides-/Average-Weight-Legislation-A-quick-reference-guide/