

VE401 TERM PROJECT

---

# **VERIFICATION OF SAMPLING CRITERION IN TESTING FOR NET QUANTITY OF PREPACKAGES WITH FIXED CONTENT**

---

April 17, 2023

Project Group 26

Meng Yuqi

Tanchavalit Ekkanat

Liu Yuzhuo

Zhang Maizhe

# Contents

<b>1</b>	<b>Abstract</b>	<b>1</b>
<b>2</b>	<b>Introduction</b>	<b>2</b>
<b>3</b>	<b>Glossary</b>	<b>2</b>
<b>4</b>	<b>Criterion of Acceptable Batch and Sample; Scheme of Sampling</b>	<b>2</b>
4.1	Generally Criterion of Batch and Sample . . . . .	3
4.2	Scheme of Sampling . . . . .	4
4.2.1	Scheme of Sampling in [JJF2005] . . . . .	4
4.2.2	Scheme of Sampling in [OIML2016] . . . . .	5
<b>5</b>	<b>Hypothesis Testing</b>	<b>5</b>
5.1	Symmetry in Neyman-Pearson Decision Theorem - A Discussion . . . . .	6
5.2	Testing for $H_0$ Using Student's $T$ -Distribution . . . . .	6
5.3	Testing against $H_0$ Using Noncentral $T$ -Distribution . . . . .	6
5.4	Insight on Hypothesis Testing . . . . .	6
<b>6</b>	<b>References</b>	<b>7</b>

# 1 Abstract

## 2 Introduction

As prepackaged food production becomes prevalent nowadays, inspection on whether the product contains the same amount of content as labeled becomes an important approach to protect the interests of the consumers. However, there are unavoidable errors in the production, which makes it impossible to ensure that every package produced contains content at least as much as what is specified in the label. It would in turn be unfair to the manufacturers as on average they would have to fill more content when packaging the products. Therefore, a statistically reasonable criterion is called for to resolve this dilemma, so that a general agreement is reached upon on how much error is allowable for a certain batch of production; and a standard is derived from the criterion to specify how the random sampling and inspection should be carried out to determine whether the production satisfies the requirement or not.

In our project, we refer to two specific criterion that is published by Chinese government [JJF2005] and International Organization of Legal Metrology [OIML2016], respectively, to explain how the "statistically reasonable" is established, and verify whether the sampling plan proposed in the two criteria aligns with the requirements that are specified in their materials. We would first look into the [JJF2005] first and cross-validate using [OIML2016]. A comparison between the two criteria will be made; and an empirical sample carried out by ourselves will be demonstrated for supplement.

## 3 Glossary

In order to avoid confusion in the subsequent elaboration on the contents, we first specify the meaning of some terminologies; and throughout the report we will align to these expressions.

- **Batch:** the assembly of all products produced whose being accepted or not depends on the samples taken from the batch.
- **Sample:** a fraction of products that are taken from the batch for examination.
- **Package:** an individual in the sample or in the batch.

## 4 Criterion of Acceptable Batch and Sample; Scheme of Sampling

Before we dig into the technical details, we first summarize the requirements that are set in both standards ([JJF2005] and [OIML2016]), on both the whole batch and on samples obtained. Inter-

estingly, although they share the same standard on batches, the criteria that are used for sample testing are quite different.

#### 4.1 Generally Criterion of Batch and Sample

First we present two definitions that will be used intensively in subsequent discussions, which are defined in [OIML2016]:

**Definition 4.1** ( $T_1$  Error,  $T_2$  error). *For a given standard  $Q$  with the applicable tolerable deficiency  $T$ , a sample  $Q_i$  is said to be of  $T_1$  error if*

$$Q - 2T \leq Q_i < Q - T$$

*and is said to be of  $T_2$  error if*

$$Q_i < Q - 2T$$

According to both criteria, the samples from a certain batch should satisfy that

- (A.1) The sample mean is no less than the labeled nominal amount.
- (A.2) The proportion of samples with nominal amount less than the required amount (labeled amount minus the tolerable deficiency) is less than 2.5%.
- (A.3) There are no  $T_2$  errors in the samples.

and the inspection process should fulfill the requirement to detect defective batches<sup>1</sup>:

- (B.1) If a shipment is correctly manufactured, i.e. with  $\mu \geq Q$  where  $Q$  is the labeled amount, then the probability of this being rejected is less than 0.5%.
- (B.2) If the sample mean is less than  $Q - 0.74s$ , then 90% of the time we can spot it out (and reject that).
- (B.3) If the proportion of  $T_1$  error is no more than 2.5%, then the probability of it being rejected is no more than 5%.
- (B.4) If the proportion of  $T_1$  and  $T_2$  error combined is 9%, then 90% of the time we can spot it out (and reject that).

---

<sup>1</sup>The sequence of presenting the rules is slightly changed so that they are classified in terms of statistic examined, instead of type of errors.

with (B.1,2) specifying the requirement of the mean and (B.3,4) specifying the requirement of the variance (as the restriction of Type I and II Errors are a restraining factor of the variance of the production). All these rules will be elaborated on in the next section.

## 4.2 Scheme of Sampling

Although the organizing ideas in both criterion are the same, the suggestions they provided in obtaining the samples are quite different. The major difference is in how the sample size  $n$  should relate to the size of the population (size of batch)  $N$ .

### 4.2.1 Scheme of Sampling in [JJF2005]

In [JJF2005], the choice of sample size  $n$  should be selected according to the following table<sup>2</sup>:

$N$ (Batch Size)	$n$ (Sample Size)	$\text{Log}(N)$	$\text{Log}^2(N)$	$n/\text{Log}^2(N)$
1-10	$N$	N/A	N/A	N/A
11-50	10	2.40	5.75	1.74
51-99	13	3.93	15.46	0.84
100-500	50	4.61	21.21	2.36
501-3200	80	6.22	38.65	2.07
>3200	125	8.07	65.14	1.92

Table 1: Designated Sample Size  $n$  Given Batch Size  $N$

There is a general trend in the selection of  $n$ : the larger  $N$  is, the smaller the incremental step of  $n$  takes. This much resembles the behavior of a logarithm-like function; therefore of our interest we list statistics of  $\text{Log}(N)$  as well. And it actually should: heuristically, the sample size  $n$  should be able to organize changes in batch size  $N$  faster than linearly, and logarithm seem to be a good fit in this case.

Although we have to admit that logarithm is a quite brutal, unverified and not so approximate modelling of the data, it could yield some insights. Specifically, we can see all rows applicable to doing the division  $n/\text{Log}^2(N)$  except for the third one yields a results of  $2.0 \pm 0.4$ . In later parts we will verify whether the criterion listed here makes full sense.

<sup>2</sup>Entry “ $\text{Log}(N)$ ” and “ $\text{Log}^2(N)$ ” takes the lower bound of  $N$  in each row

#### 4.2.2 Scheme of Sampling in [OIML2016]

We similarly present the table specified in [OIML2016] according to which the sample size should be chosen:

$N$ (Batch Size)	$n$ (Sample Size)	$\text{Log}(N)$	$\text{Log}^2(N)$	$n/\text{Log}^2(N)$
1-20	$N$	N/A	N/A	N/A
21-40	32	3.04	9.27	3.45
41-60	35	3.71	13.79	2.54
61-80	47	4.11	16.90	2.78
81-100	49	4.39	19.31	2.54
100-200	64	4.61	21.21	3.01
200-300	67	5.30	28.07	2.39
300-400	81	5.70	32.53	2.49
400-500	81	5.99	35.90	2.27
>500	98	6.21	38.62	2.54

Table 2: Designated Sample Size  $n$  Given Batch Size  $N$

The statistic generally falls into a centered region of  $2.9 \pm 0.6$ . The anomaly 2.27 appears at the second-to-bottommost row, but the same  $n$  appearing in consecutive columns is also a bit weird. We can vaguely sense that this table is less likely to be wrong than [JJF2005]; but further investigations are required, which we will present in juxtaposition with the that of the previous table.

## 5 Hypothesis Testing

The statistic foundation of formulating the probability of various situations is hypothesis testing. In order to rigorously conduct discussions on the probability of accepting or rejecting a certain package, we need to formalize them using Neyman-Pearson Decision Theory [Ho2023], accept or reject according to the sample falling into critical region or not, and comment on Type II Error and the power of the test. We will resolve these problems in this section.

Besides explaining the source of the statistics listed in the table, we would also like to look into the symmetry in Neyman-Pearson Decision Theorem. The criterion gives a test using central  $T$ -distribution where  $\mu = \mu_0$  where  $\mu$  is the sample mean and  $\mu_0$  is the labeled amount is the boundary case. Could we perform in the other way, using the boundary value in the alternative hypothesis to carry out this testing? The above question will be discussed in depth in the following subsections.

**5.1 Symmetry in Neyman-Pearson Decision Theorem - A Discussion**

**5.2 Testing for  $H_0$  Using Student's  $T$ -Distribution**

**5.3 Testing against  $H_0$  Using Noncentral  $T$ -Distribution**

**5.4 Insight on Hypothesis Testing**



## 6 References

- [Ho2023] Horst Hohberger. Probabilistic Methods in Engineering (lecture slides). University of Michigan - Shanghai Jiao Tong University Joint Institute. Spring 2023. (An earlier version can be found at [http://ece401.ji.sjtu.edu.cn/ewExternalFiles/ve401\\_all\\_lecture\\_slides.pdf](http://ece401.ji.sjtu.edu.cn/ewExternalFiles/ve401_all_lecture_slides.pdf))
- [JJF2005] AQSIQ/MTC1. Rules of metrological testing for net quantities of products in prepackages with fixed content. Technical Report JJF 1070-2005, Chinese Metrology Press, 2005. Available from <http://www.gdifi.org.cn/spbz/437.jhtml> [Online; accessed 17-April-2023].
- [OIML2016] Quantity of product in prepackages. Technical Report OIML R 87, International Organization of Legal Metrology (OIML), 2016. Available from [https://www.oiml.org/en/files/pdf\\_r/r087-e16.pdf](https://www.oiml.org/en/files/pdf_r/r087-e16.pdf) [Online; accessed 25-March-2023].