Sampling considerations within Market Surveillance actions

Nikola Tuneski, Ph.D.
Department of Mathematics and Computer Science
Faculty of Mechanical Engineering
Skopje, Macedonia

Sampling considerations within Market Surveillance actions

Content:

I. Present situation
II. Ideas for improvement

Sampling considerations within Market Surveillance actions

I. Present situation

- In general ISO 2859-1 (Sampling procedures for inspection by attributes) is used.
- ISO 2859-1 was developed for production quality control.

Sampling considerations within Market Surveillance actions

I.1 Sample size by ISO 2859-1:

- Depends on the size of the lot and the inspection level.

- Disadvantage: Brings large sample sizes for big lots and rises the cost of the inspection.

Sampling considerations within Market Surveillance actions

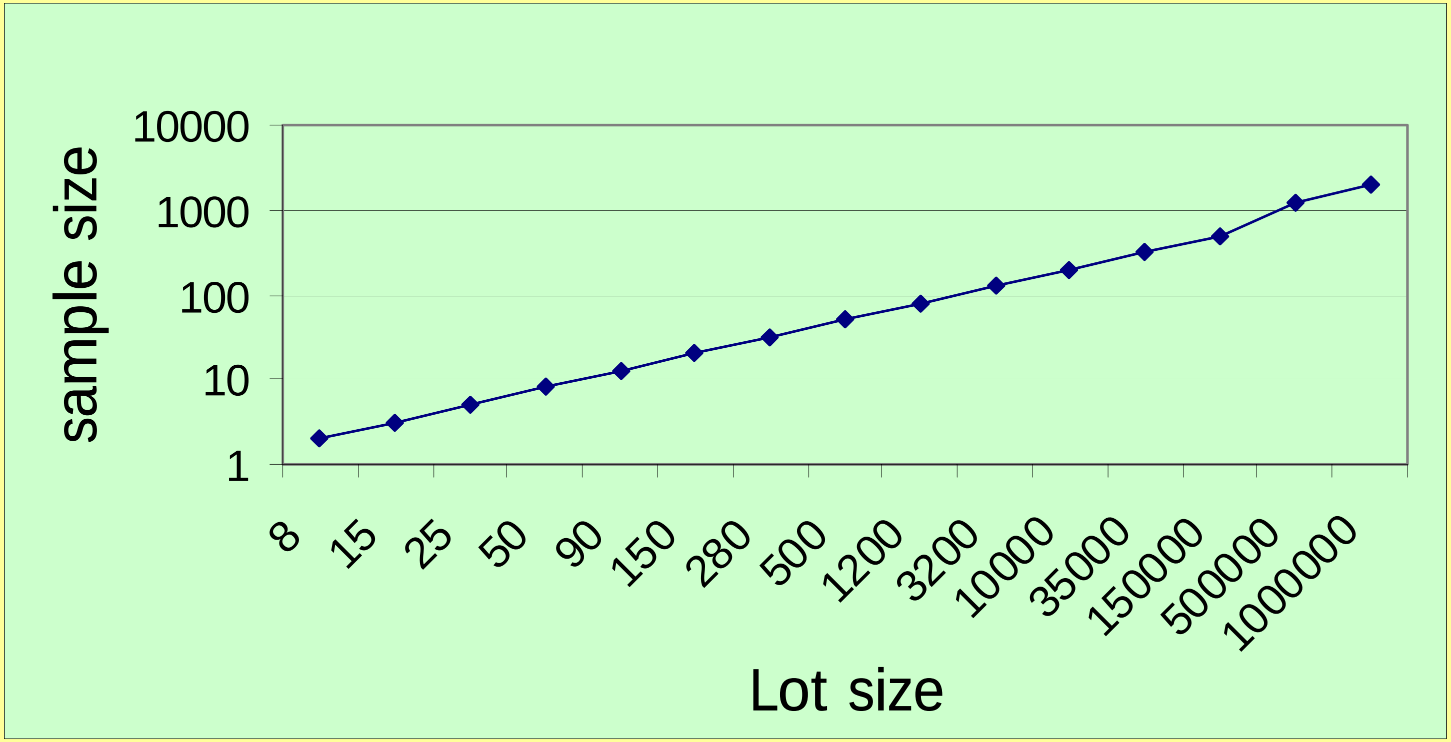
For inspection level II:

Lot size	Sample size
2 to 8	2
9 to 15	3
16 to 25	5
26 to 50	8
51 to 90	13
91 to 150	20
151 to 280	32
281 to 500	50

Lot size	Sample size
501 to 1200	80
1201 to 3200	125
3201 to 10000	200
10001 to 35000	315
35001 to 150000	500
150001 to 500000	1250
over 500001	2000

Sampling considerations within Market Surveillance actions

In logarithmic scaling:



Sampling considerations within Market Surveillance actions

II. Ideas for improvement

- Binomial distribution.
- Bayesian statistics.

Sampling considerations within Market Surveillance actions

II.1 Binomial distribution

Sample size: $$n = \frac{z_{LC}^2 p(1-p)}{E^2}$$

LC is the level of confidence (0.9, 0.95,...)

z_{LC} is the z-value for level of confidence LC

p is a prior knowledge about the proportion of defective items in the lot (if not known p=0.5)

E is margin of error

Sampling considerations within Market Surveillance actions

- Reference:
  - ISO 11453 (Statistical interpretation of data - Tests and confidence intervals relating to proportions)

- Advantages:
  - Tailor made sample sizes that does not depend on the size of the lot.
  - For large lots implies smaller sample sizes than the ones obtained by ISO 2859-1.

- Disadvantage:
  - It can be used only when np and $n(1-p)$ are bigger or equal to 5.

Sampling considerations within Market Surveillance actions

II.2 Bayesian statistics

Gives excellent results when small samples are given.

The main idea behind the Bayesian statistics is that the parameters (proportion, mean, deviation,...) are viewed as random variables.

In classical statistics they are viewed as variables with fixed (unknown) value.

Sampling considerations within Market Surveillance actions

Since a parameter (denoted by p) is now a random variable, its probability distribution can be specified.

Such a distribution is called **prior distribution** (denoted by $\pi(p)$) and it usually reflects our prior beliefs about the parameter.

Sampling considerations within Market Surveillance actions

Once the experiment (inspection) is conducted and sample (x) is observed **posterior distribution** of the parameter can be obtained that collects all knowledge (prior and sample):

$$\pi(p \mid x) = \frac{f(x \mid p) \cdot \pi(p)}{m(x)}$$

$f(x|p)$ is the distribution of the population given (or with respect to) the parameter p.

$m(x)$ is the marginal distribution of x.

Sampling considerations within Market Surveillance actions

After the posterior distribution is obtained we can calculate a $100LC$% **Bayesian interval** (a,b) by solving the following two equations

$$\int_{-\infty}^{a} \pi(p \mid x)dp = \int_{b}^{+\infty} \pi(p \mid x)dp = \frac{1-LC}{2}$$

Sampling considerations within Market Surveillance actions

Example:

It is known (prior knowledge) that the percentage of defective bulbs in a lot is uniformly distributed between p_1=0% and p_2=20%. In a sample of n bulbs we find x defective.

Level of confidence LC=95%.

Sampling considerations within Market Surveillance actions

n	*x*	Bayesian		"Traditional"	
		Interval (in %)	Width (in %)	Interval (in %)	Width (in %)
3	0	(0.37, 19.29)	18.92		
	1	(2.80, 19.70)	16.90	(0, 86.68)	86.68
	2	(05.62, 19.82)	14.20	(13.32, 100)	86.68
5	0	(0.31, 19.09)	18.78		
	1	(2.48, 19.65)	17.17	(0, 55.06)	55.06
	2	(5.18, 0.1980)	14.62	(0, 82.94)	82.94
	3	(7.52, 19.86)	14.68	(17.06, 100)	82.94
10	0	(0.21, 18.27)	18.06		
	1	(1.86, 19.44)	17.58	(0, 28.59)	28.59
	2	(4.21, 19.71)	15.50	(0, 44.79)	44.79
	3	(6.48, 19.82)	13.34	(1.59, 58.4)	56.81
	4	(8.41, 19.87)	13.39	(9.64, 70.36)	60.73

Sampling considerations within Market Surveillance actions

- Advantages:
  - Better estimates for same sample size.
  - Smaller sample size and margin of error comparing with the binomial approach.

- Disadvantages:
  - "Bad" choice of the prior distribution leads to wrong conclusions.
  - There are no explicit formulas for the Bayesian interval or for the sample size.
  - Complicated calculations not suitable for practical use.

Sampling considerations within Market Surveillance actions

- Research questions:
  - Developing mathematical procedures for calculating Bayesian interval and sample size for different types of prior distributions.
  - Writing a computer program(s) that performs those calculations.
  - Creating tables and/or charts for practical usage.