Acceptance Sampling (the most famous version of product-oriented sampling inspection)

Lot N items

Sample n items

↳ "accept" lot or "reject" lot

There are both "attributes" and "variables" version.

based on counts of defects or defectives

based on measurements and summaries like $\bar{x}$, s

"Attributes" Versions (and in particular, single sampling)

X = # of defectives in sample (% defective context)

or X = total # of defects on sampled items (mean defects per unit context)

If $X \le c$ "accept" lot

If $X > c$ "reject" lot

acceptance number

Quantification of what such a regimen will do is usually made in terms of the OC (operating characteristic):

$$OC = P_a = P[X \le c].$$

This is usually thought of as a function of either "p" or ("$\frac{T}{N}$" or "λ").

<u>% defective contexts</u>

p could be the lot fraction defective $(0, \frac{1}{N}, \frac{2}{N}, \cdots, 1)$ } perspective A

or p is a constant propensity for any item to be defective } perspective B

<u>Mean Defects/Unit Contexts</u>

$\frac{T}{N}$ actual occurence rate of defects in lot } perspective A

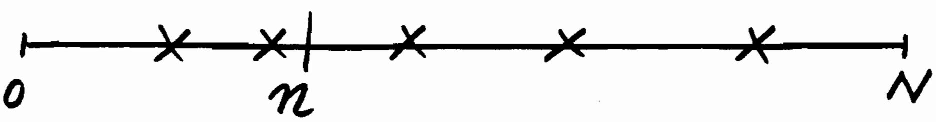
(T: total # of defects on N items)

λ mean occurence rate for a process generating flaws (Poisson process parameter) } perspective B

	B	A
% Defective	stable process model for G/D characteristics $X \sim Bin(n, p)$	Under simple random sampling with $N \cdot p$ defectives in lot $P(X=x) = \dfrac{\binom{Np}{x}\binom{N(1-p)}{n-x}}{\binom{N}{n}}$
Mean Defects/Unit	Poisson process model with rate λ $X \sim Poisson(n\lambda)$	Poisson process model (with constant rate) conditional on T $X \sim Bin(T, \frac{n}{N})$

Why is the dsn in lower right correct?

Basic probability fact: If events are occurring according to a Poisson process and I have T in a period N, then conditional on T the locations are iid Uniform $[0, N]$.



If $T=5$, conditional on this locations are uniform on the interval.

It turns out that in many cases of (N, n, c) acceptance probabilities calculated in any of the above four ways are similar.

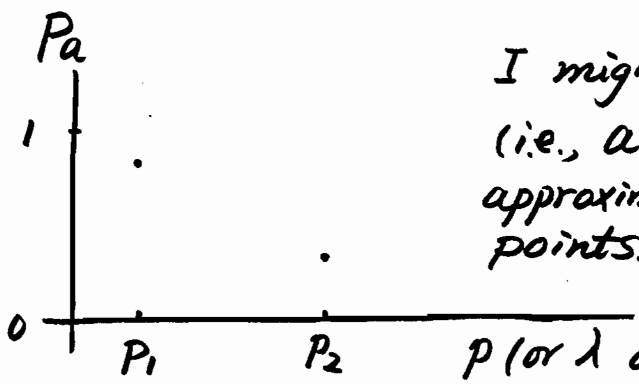
— Note Figures 8.1 and 8.2 of V&J in this regard. (page 451)

<u>Design</u> of an attributes acceptance sampling plan (single sampling) amounts to choice of n, c — we'll talk about two approaches:

"Classical two-point method" and

"Economic / Decision Theoretic" choice

<u>Classical two-point method</u>:



I might hope to choose (n, c) (i.e., an OC curve) running approximately through those points.

Devices to do this (at least approximately):

1) Formulas (8.17) & (8.18) of V&J based on a normal approximation to dsn of X. (p. 455)

2) The method in (8.19) through (8.21) based on a Poisson $(n\lambda)$ dsn for X. (p. 457)

Economic / Decision Theoretic Choice:

Idea: Use enough of a probability and cost structure to find expected cost as function of (n, c) and optimize.

We'll think through (Barlow's formulation of) Deming's Inspection Problem.

Lot of N items

K_1 = cost of inspecting a single item

K_2 = cost of later grief caused by any defective item that goes uninspected

Suppose inspection is perfect.

This is the Vander Wiel / Vardeman cost scenario where

$K_I = K_1$, $K_{DF} = 0$, and $K_{DU} = K_2$,

and $W_G = W_D = 0$.

Do "single sampling with 'rectification'."

Let X = # of defectives in sample.

If $X \le c$, accept lot;
(replace any D's in sample
but do no more inspection)

if $X > c$, reject lot.
(do 100% inspection, replacing all D's)

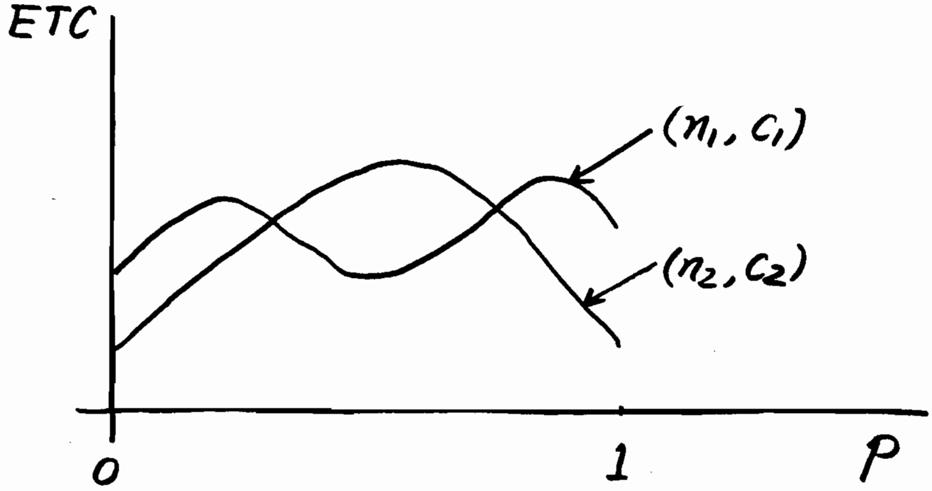
(Assume that "replacements" are
guaranteed G and "free.")

<u>Probability Model</u>: perspective B / stable
process model with

p = <u>constant</u>, propensity for
each item to be D

for the lot under consideration

$$ETC(n, c, p) = nK_1 + P_a(N-n)pK_2$$
$$+ (1-P_a)(N-n)K_1$$
$$= K_1 N\left[1 + P_a\left(1-\frac{n}{N}\right)\cdot\left(p\frac{K_2}{K_1} - 1\right)\right].$$



- How to choose between (n, c) pairs?

The answer depends upon what I'm willing to input about p.

<u>Possibility 1</u>: p known

(leads to inspecting "All or None")

<u>If $p > \frac{K_1}{K_2}$</u> (so $p\frac{K_2}{K_1} - 1 > 0$)

an optimal policy will be one where $P_a = 0$ or $1 - \frac{n}{N} = 0$, i.e., we should <u>inspect "All."</u>

<u>If $p < \frac{K_1}{K_2}$</u> (so $p\frac{K_2}{K_1} - 1 < 0$)

an optimal policy will be one where $P_a = 1$ and $1 - \frac{n}{N} = 1$, i.e., we should <u>inspect "None."</u>

Possibility 2: p not known but I am willing to describe it with a dsn G on $[0,1]$.

This might make sense from two perspectives:

1) I'm a Bayesian with a "prior" dsn for p (G is the "prior").
2) I'm modeling "process variation" in p (G is the "process dsn").

Once I settle on a G, I can average

$$ETC(n, c, p)$$

over that dsn, thereby eliminating p.

Having done that, I can then compare not functions of p, but rather numbers for various (n, c) combinations to get a best one.

"Sample" on Professor Vardeman's Stat 531 web page will do this for you (do the optimization) for the case where G is Beta dsn (allowing a slightly more general cost structure).

Let's consider the mathematics of this (up to the point where it gets messy).

Begin by picking an optimal c for each sample size

$$c_G^{opt}(n).$$

Note that if a sample of n results in X defectives then the conditional expected cost is

With no additional inspection: $nK_1 + (N-n)K_2 E_G[p|X]$

If I inspect the remainder: NK_1

So I should do no more inspection if

$$nK_1 + (N-n)K_2 E_G[p|X] < NK_1,$$

i.e., if

$$E_G[p|X] < \frac{K_1}{K_2}.$$

The prescription here simply replaces p with $E_G[p|X]$ in the "known p" analysis.

An optimal choice of c is therefore

$$c_G^{opt}(n) = \max\left\{x \,\middle|\, E_G[p|X=x] \le \frac{K_1}{K_2}\right\}.$$

By the way, it turns out that $E_G[p|X=x]$ is monotone nondecreasing in x. This is a consequence of the monotone likelihood ratio property of the binomial dsn.

$$\left.\begin{array}{l} X|p \sim \text{Binomial}(n,p) \\ p \sim G \end{array}\right\} \Rightarrow \text{a joint dsn for } (X,p)$$

$\Rightarrow$ conditional dsns for $p|X=x$.

Look at the means of those dsns and find the largest x so that the mean is still below K_1/K_2. That is $c_G^{opt}(n)$.

Then what?

Compute and compare

$$E_G[ETC(n, c_G^{opt}(n), p)]$$

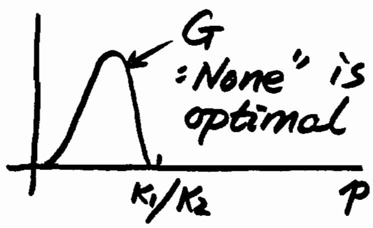
for various n. The best n is the one with the smallest corresponding value.

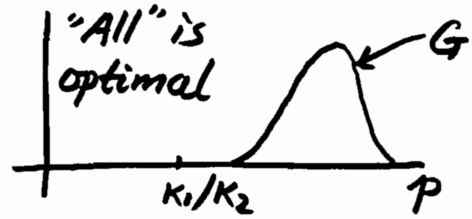
In this model depending upon the nature of G it can turn out that neither "All" $(n=N)$ nor "None" is optimal.

<u>Example</u>. $K_1 = 1$, $K_2 = 1000$; G: $P[p=0] = \frac{1}{2} = P[p=1]$.

"clearly" $n \equiv 1$ is optimal.

<u>Example</u>. If G puts all its mass on a single side of $\frac{K_1}{K_2}$ then "clearly" either "All" or "None" is optimal.





So the "interesting" problems are ones where G put probability on both sides of K_1/K_2.

<u>Example</u>. G the Beta dsn (parameters α, β).
(This is the most tractable version of this analysis.)

Here $\quad E_G[p \mid X = x] = \dfrac{\alpha + x}{\alpha + \beta + n}$.

So $c_G^{opt}(n)$ is the largest x so that

$$\frac{\alpha + x}{\alpha + \beta + n} \le \frac{K_1}{K_2},$$

i.e., $c_G^{opt}(n) = \lfloor \frac{K_1}{K_2}(\alpha + \beta + n) - \alpha \rfloor$ greatest integer function

$$= \lfloor \tfrac{K_1}{K_2} n - \alpha + \tfrac{K_1}{K_2}(\alpha + \beta) \rfloor$$

$$\approx \tfrac{K_1}{K_2} n \quad \text{for large } n.$$

The optimal n is found by minimizing (over choices of n)

$$E_G\left[ETC\left(n, c_G^{opt}(n), p\right)\right]$$

$$= \int_0^1 ETC\left(n, c_G^{opt}(n), p\right) \cdot \frac{1}{B(\alpha,\beta)} p^{\alpha-1}(1-p)^{\beta-1} dp.$$

This boils down to minimizing

$$\left(1-\frac{n}{N}\right) \cdot \sum_{x=0}^{c_G^{opt}(n)} \binom{n}{x} \cdot \int_0^1 p^x (1-p)^{n-x} \cdot \left(p\frac{K_2}{K_1}-1\right) \cdot p^{\alpha-1}(1-p)^{\beta-1} dp$$

as a function of n.

As a matter of fact, "Sample" does something else (equivalent).

This "Deming Inspection Problem" formulation is not the most general possible, but the solutions one gets are typical of what one finds doing economic choice of inspection plans.