Statistical Bulletin
Reliability & Variation Research

DETROIT RESEARCH INSTITUTE
BOX 1163 NORTHLAND CENTER STATION • SOUTHFIELD, MICHIGAN 48075
(313) 353-4556

Leonard G. Johnson, EDITOR

Vol. 3

Bulletin No. 1

April, 1973

UPPER CONTROL LIMITS
FOR AUDITING AUTOMOTIVE EMISSION RATES

BASIC RULE OF QUALITY CONTROL

RULE: No sample parameter should be more than three standard deviations (of the sample parameter) above the population value to be maintained for that parameter.

This implies that there is a maximum value permitted for a sample parameter. As long as this maximum permissible value is not exceeded, it can be assumed that the process or product involved has not gone out of control at the high end. This maximum permissible value is known as the Upper Control Limit for the sample parameter under consideration. The formula for the upper control limit (UCL) is

$$UCL = \text{Population Value of Parameter} + 3 \cdot \sigma_{\text{Sample Parameter}}$$

Where,

$$\sigma_{\text{Sample Parameter}} = \begin{array}{c}\text{Standard Deviation of the Sample Parameter} \\ \text{For the Size of Samples Assumed}\end{array}$$

In the control of emission levels for a particular pollutant (say hydrocarbons), we have three population parameters to be maintained. These are:

1. The Population Mean, μ, for the emission level in grams per mile.
2. The Population Standard Deviation, σ, for the emission level in grams per mile.
3. The Population Skewness, α_3, of the distribution of the emission level. This defines the shape of the distribution of emission rates.

To keep the three parameters under control, we must define their Upper Control Limits for whatever sample size is employed in an audit program. The sample size employed is dictated by convenience rather than by any statistical requirement. The upper control limits on sample parameters are statistical requirements which vary with sample size.

I: The Upper Control Limit For a Sample Mean

In applying the general formula:

$$UCL = \text{Population Value of Parameter} + 3\sigma_{\text{Sample Parameter}}$$

We have:

$$\text{Population Value of Parameter} = \text{Population Mean} = \mu$$

$$\sigma_{\text{Sample Parameter}} = \sigma_{\text{Sample Mean}} = \frac{\sigma}{\sqrt{N}}$$

Where σ = Population Standard Deviation

and N = Sample Size Employed in Auditing

Hence, for the parameter called the mean, we have . . .

$$UCL = \mu + \frac{3\sigma}{\sqrt{N}}$$

Formula for Upper Control Limit of a Sample Mean

II: The Upper Control Limit for A Sample Standard Deviation

In applying the general formula

$$UCL = \text{Population Value of Parameter} + 3\sigma_{\text{Sample Parameter}}$$

We have:

$$\text{Population Value of Parameter} = \text{Population Standard Deviation} = \sigma$$

$$\sigma_{\text{Sample Parameter}} = \sigma_{\text{Sample Std. Dev.}} = \frac{\sigma}{\sqrt{2N}}\sqrt{1+\frac{3}{4}\alpha_3^2}$$

Where α_3 = Population Skewness

and N = Sample Size Employed in Auditing

Hence, for the parameter known as the Standard Deviation, we have . . .

$$\text{UCL} = \sigma + \frac{3\sigma}{\sqrt{2N}}\sqrt{1+\frac{3}{4}\alpha_3^2}$$

Formula for Upper Control Limit of a Sample Standard Deviation

III: The Upper Control Limit for a Sample Skewness

In applying the general formula

$$\text{UCL} = \text{Population Value of Parameter} + 3\,\sigma_{\text{Sample Parameter}}$$

We have:

Population Value of Parameter = Population Skewness = α_3

$$\sigma_{\text{Sample Parameter}} = \sigma_{\text{Sample Skewness}} = \sqrt{\frac{6}{N}}$$

Where N = Sample Size Employed in Auditing.

Hence, for the parameter known as the Skewness, we have . . .

$$\text{UCL} = \alpha_3 + 3\sqrt{\frac{6}{N}}$$

Formula for Upper Control Limit of a Sample Skewness.

HYPOTHETICAL NUMERICAL EXAMPLE

For maintaining a pollutant's emission parameters of

$$\mu = 1.87 \text{ grams/mile}$$

$$\sigma = 0.78 \text{ grams/mile}$$

$$\alpha_3 = 1.24$$

Suppose these population parameters have been established by testing several thousand vehicles.

Using the formulas just developed, we obtain the following table of upper control limits for various sample sizes:

Sample Size	UCL (Sample Mean)	UCL (Sample Std. Dev.)	UCL (Sample Skewness)
10	2.61	1.55	3.56
20	2.39	1.32	2.88
30	2.30	1.22	2.58
40	2.24	1.16	2.40
50	2.20	1.12	2.28
60	2.17	1.09	2.19
70	2.15	1.07	2.12
80	2.13	1.05	2.06
90	2.12	1.04	2.01
100	2.10	1.02	1.97
200	2.04	0.95	1.76
300	2.01	0.92	1.66
400	1.99	0.90	1.61
500	1.97	0.89	1.57
600	1.97	0.88	1.54
700	1.96	0.87	1.52
800	1.95	0.86	1.50
900	1.95	0.86	1.48
1000	1.94	0.86	1.47

REJECTION OF OUTLIERS

When should an automobile be considered an outlier as far as emmission rates are concerned? This type of question has always been the main concern in any quality control program. Over the years it has become an adopted custom in quality control circles to consider any item within the so-called 3 sigma limits about the mean as passing, and any item outside of these limits has come to be defined as an outlier. Schematically, this concept can be represented as follows:

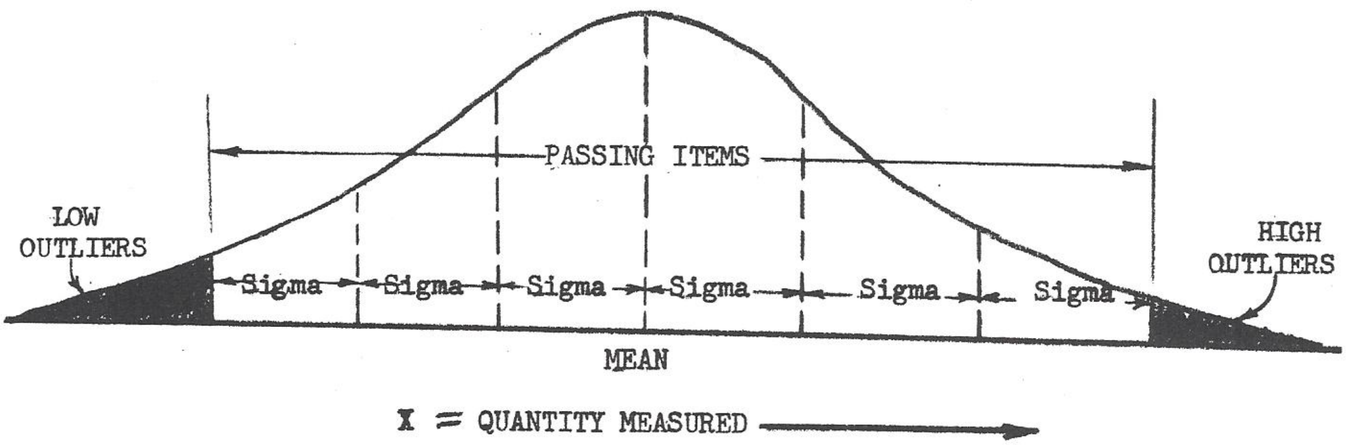


FIGURE 1

In FIGURE 1, which is a frequency distribution for some measured property of an item, we see that the passing items have measurements lying between (MEAN - 3 Sigma) and (MEAN + 3 Sigma).

Low outliers have measurements less than (MEAN - 3 Sigma).

High outliers have measurements greater than (MEAN + 3 Sigma).

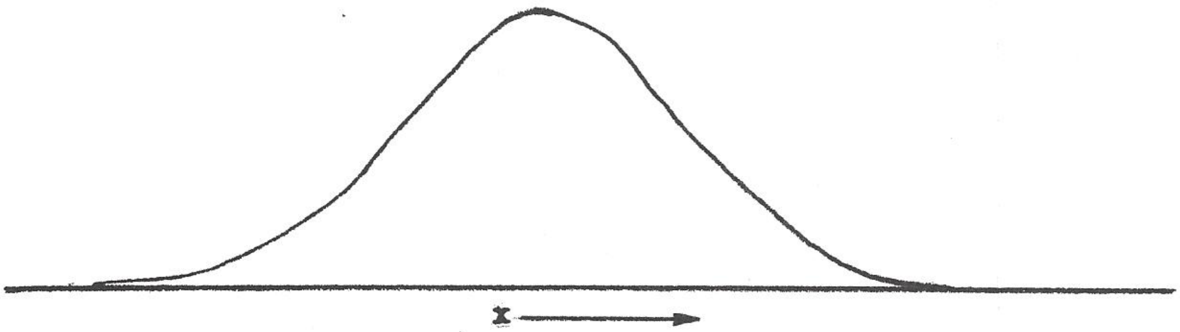


FIGURE 2 —NORMAL DISTRIBUTION

(Skewness = 0)

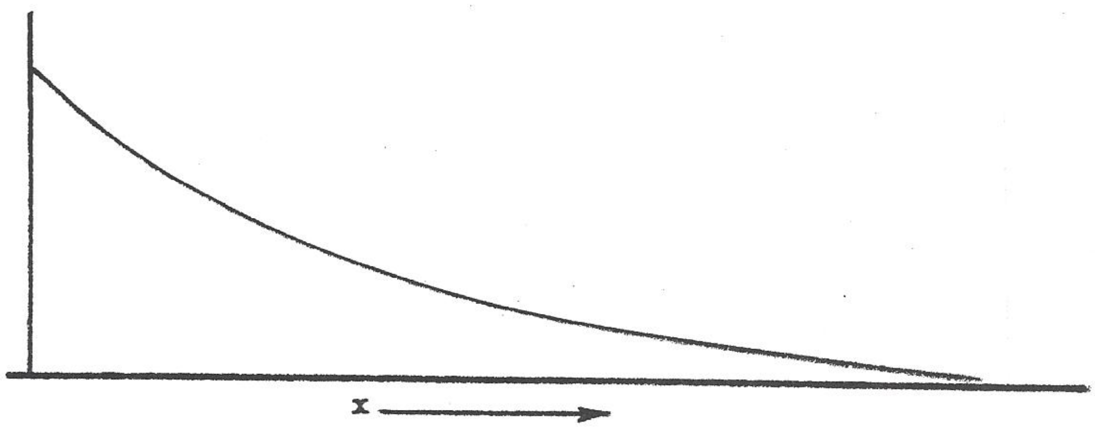


FIGURE 3 — EXPONENTIAL DISTRIBUTION

(Skewness = 2)

Between these two cases there are intermediate skewnesses (such as 1), which have the appearance of FIGURE 4 below.

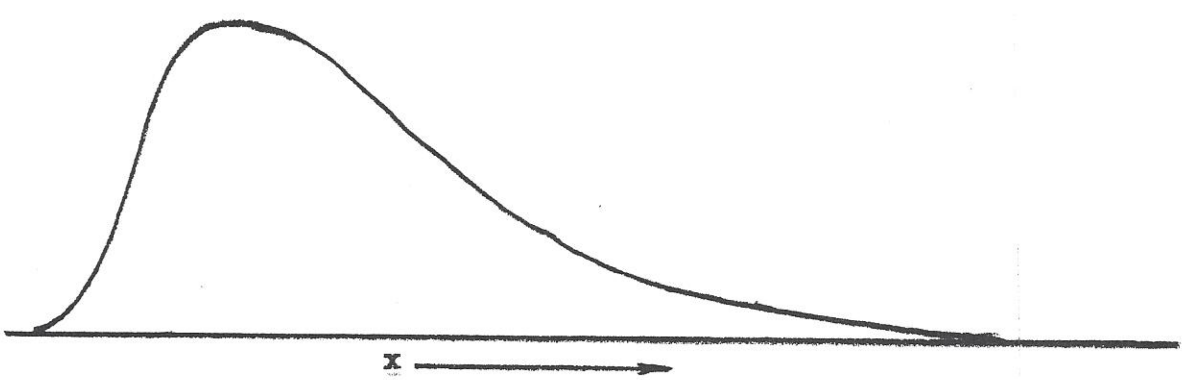


FIGURE 4 —BETWEEN EXPONENTIAL & NORMAL

(Skewness between 0 and 2)

In the problem of emmission rates of automobiles we are really only concerned with high outliers, since very low emission rates are not ecologically bad. Now we ask the following question:

"Suppose an automobile has an emission rate exceeding (MEAN + 3 Sigma) for a given population of automobiles, what are the chances that it really is not an outlier (i.e., alien to the population) even though by the 3 Sigma concept it is arbitrarily defined to be an alien to the population?"

The answer to this question depends very much on the shape of the frequency distribution shown in FIGURE 1, and, more specifically, it depends upon the nature of the right hand tail of the distribution beyond (MEAN + 3 Sigma).

Generally speaking, the shape of a distribution is defined by a shape parameter known as the skewness, which is nothing more than the sum of all third moments about the mean divided by the cube of the standard deviation (Sigma). We are here considering only unimodal distributions (those having a single peak, or mode).

The so-called NORMAL (GAUSSIAN) DISTRIBUTION has a skewness of zero (FIGURE 2).

The EXPONENTIAL DISTRIBUTION (with peak at extreme left) has skewness 2 (FIGURE 3).

So far we have considered only positive skewness. It is also possible to have negative skewness. For example, a skewness of -2 would have the appearness of FIGURE 5.

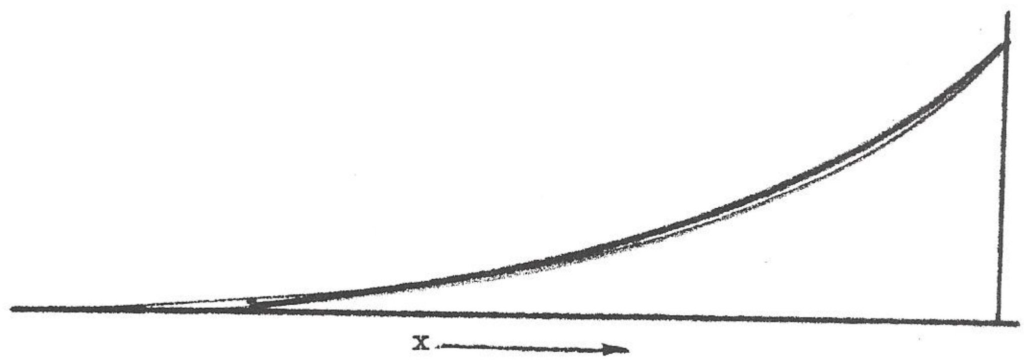


FIGURE 5 — RISING EXPONENTIAL FREQUENCY CURVE

(Skewness = -2)

We see that FIGURE 5 is a left to right mirror image of FIGURE 3.

Likewise, a left to right mirror image of FIGURE 4 would give a distribution with a skewness between 0 and -2, as shown in FIGURE 6.

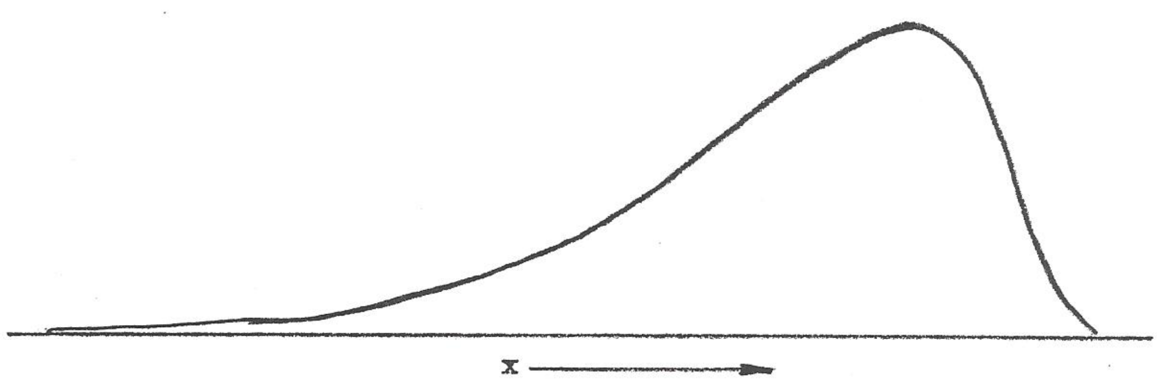


FIGURE 6 — BETWEEN RISING EXPONENTIAL & NORMAL

(Skewness between 0 and -2)

Obviously, the area of the right hand tail of a distribution beyond (MEAN + 3 Sigma) varies with the skewness. The following table shows this variation:

TABLE 1

SKEWNESS	AREA OF TAIL BEYOND (MEAN + 3 Sigma)	
0	.00135	1/741
.25	.00317	1/315
.50	.00543	1/184
.75	.00788	1/127
1.00	.01034	1/97
1.25	.01200	1/83
1.50	.0141	1/71
1.75	.0162	1/62
2.00	.0183	1/55

The areas in the right hand column of TABLE I are the probabilities that a defined outlier (beyond MEAN + 3 Sigma) is still a member of the population from which it has been cast out by the 3 Sigma concept. It can be seen that when the population is NORMAL (Zero skewness) the probability that an outcast is still a member is 1 chance in 741. For this case of zero skewness the probability is the smallest, but for increasing skewness (in the positive direction) the probability that an outcast is a member increases, until at a skewness of 2 it is 1 chance in 55.

Thus, it can be seen that even for the extreme skewness of 2 there is less than 1 chance in 50 that an item which has been rejected because of falling above (MEAN + 3 Sigma) is really a member of the population. In other words, we are at least 98% confident that the rejected item belongs to some other population, and hence should be considered an alien to the population under consideration.

For this reason the quality control custom of passing only items within ± 3 Sigma of the MEAN is a very reasonable practice. We should adopt this same practice with regard to emission levels, and reject all cars whose emission rates exceed (MEAN + 3 Sigma).

The UPPER CONTROL LIMIT as a function of MILEAGE will then be

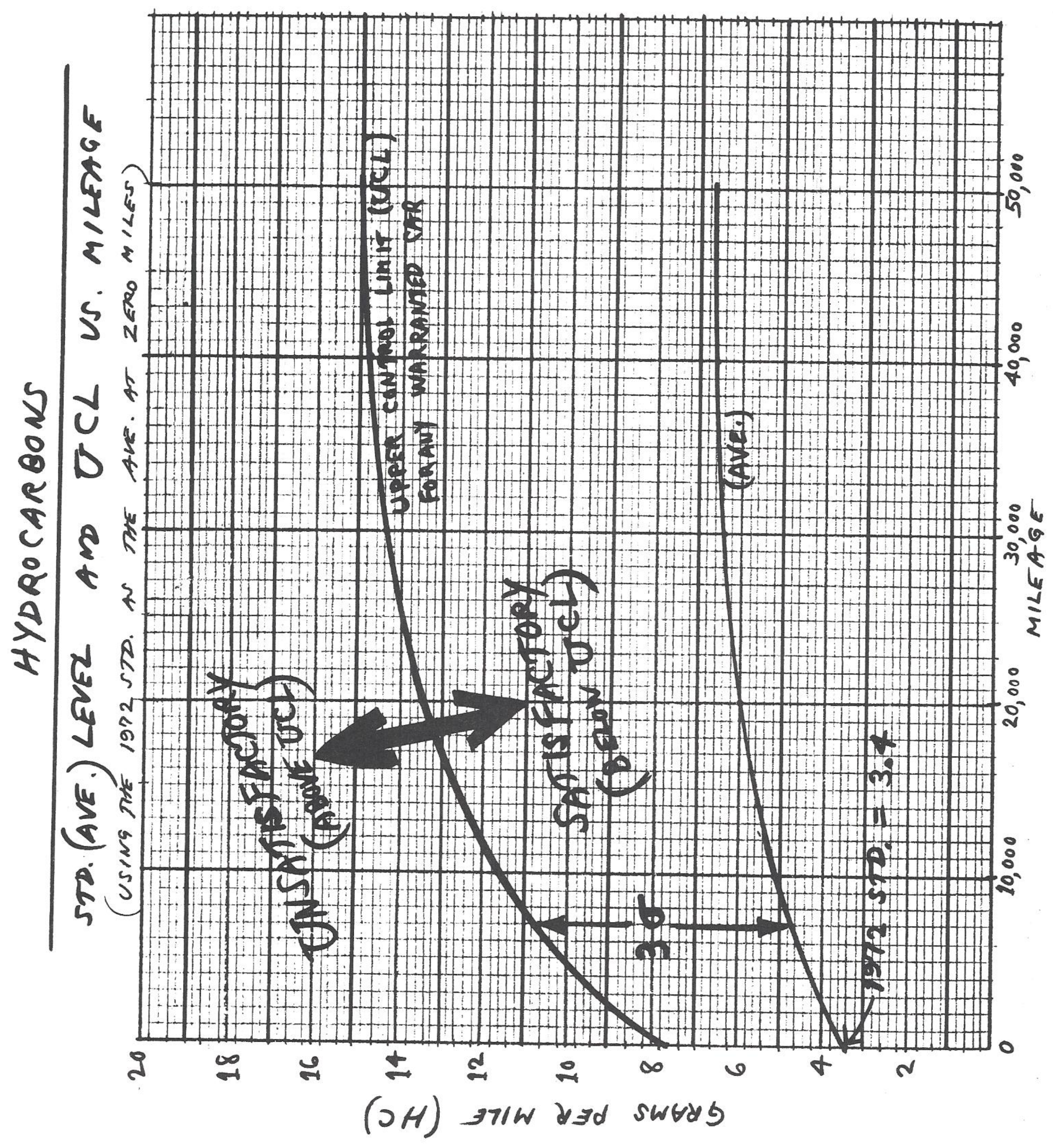
MEAN (at that mileage) + 3 SIGMA (at that mileage).

An example is shown graphically in FIGURE A .

In FIGURE A we take the average at zero miles to be the GOVERNMENT STANDARD.

The growth rate with mileage would be determined from actual test data.

FIGURE A


HYDROCARBONS
STD. (AVE.) LEVEL AND UCL VS. MILEAGE
(USING THE 1972 STD. AS THE AVE. AT ZERO MILES)
UPPER CONTROL LIMIT (UCL) FOR ANY WARRANTED CAR
(AVE.)
UNSATISFACTORY (ABOVE UCL)
SATISFACTORY (BELOW UCL)
3σ
1972 STD. = 3.4
0
10,000
20,000
30,000
40,000
50,000
MILEAGE
20
18
16
14
12
10
8
6
4
2
GRAMS PER MILE (HC)