Reliability Improvement in Automotive Engineering

TP Davis
Ford Motor Company

1999 Spring Research Conference on Statistics in Industry and Technology
Minneapolis, MN; 2-4 June, 1999.

Preliminaries

Reliability improvement is important in automotive engineering -

- brand image of the manufacturer
- increasing customer expectations
- residual value of the vehicle upon re-sale
- minimize warranty costs

There are two types of failure -

hard failures (product function ceases), and
soft failures (degraded performance).

Failures are caused by

mistakes, and
lack of robustness.

Eliminating mistakes is primarily a matter of vigilance.

Improving robustness requires ***statistical engineering***.

"Reliability is failure mode avoidance"
Don Clausing - MIT

Field Data

Field data in automotive engineering is not complete.

Warranty data (costs and repairs) - is available for every car within the warranty period; however, the data systems are primarily designed to refund dealers the repair and labour cost.

Survey data - based on sampling vehicle owners. Subject to sampling errors, but can give clues to failure modes not seen in warranty (usually soft failures with no diagnosis and no fix).

Fleet data - usually a number of pre- and early production vehicles put into the field with high mileage accumulators, for early warning of any problems missed in the engineering development.

Field Reliability

Reliability of vehicles in the field is generally impossible to quantify, because

- almost all cars are lost to follow up at high times and mileages in service;

- there are poor records of cars that do fail;

- there are many different unknown duty cycles and stresses in the field as a consequence of the number of units in the field in a variety of different markets;

- competing risks means a reliability measure for the total vehicle is difficult to formalize (better to understand reliability at the component level.);

These problems are not too dissimilar to those faced in reliability (survival) studies in medical applications.

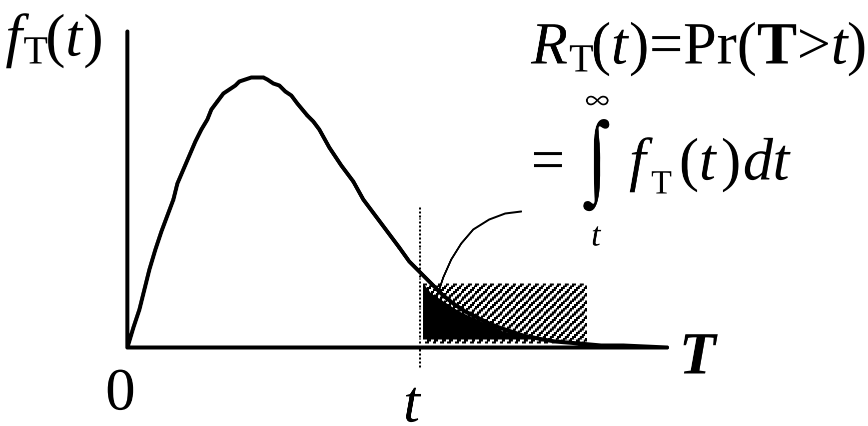
Definitions

We denote the life variable by T. The life variable can be Time, Mileage, Cycles, or something else. We use "time" as a default.

The reliability in the field at time $T=t$ is given by

$$R_T(t) = \Pr(T > t) = \int_t^{\infty} f_T(t)dt$$

where $f_T(t)$ is the probability density function for T. Of course for a given t, we want this reliability to be as large as possible (failure mode avoidance).



It is better to think of $R_T(t)$ as a *marginal* reliability because it depends on other variables such as useage stresses (denote the stresses by S).

Probably the most important function to describe data in reliability studies is the *hazard function*, denoted by $h(t)$. It describes the probability of instant failure at t, *conditional* on having survived until t. The hazard can be expressed as

$$h(t) = \frac{f(t)}{R(t)} = \frac{f(t)}{\int_t^{\infty} f(t)dt}.$$

Hence the hazard function uniquely defines the probability density function, and vice-versa. Note that $h(t)$ is strictly positive. If $h(t)$ increases with t, then this means an ***increasing failure rate*** *(IFR)*; if $h(t)$ is constant, this is ***constant failure rate*** *(CFR)*; and if $h(t)$ decreases with t, this is a ***decreasing failure rate*** *(DFR)*.

Integrating $h(t)$, we get

$$H(t) = \int_0^t h(t)dt = \int_0^t \frac{f(t)}{1-F(t)}dt = [-\ln\{1-F(t)\}]_0^t$$

$$= -\ln[1-F(t)] = -\ln[R(t)].$$

H(t) is called the *cumulative hazard function*. The *gradient* of $H(t)$ at t is $h(t)$. The above expression leads to the important result $R(t) = \exp[-H(t)]$.

We can define **conditional reliability**, i.e. reliability *conditional* on a *particular* S-value, s say, as

$$R_{T|S}(t|s) = \Pr(T > t|S = s) = \int_{t}^{\infty} f_{T|S}(t|s)dt,$$

where $f_{T|S}(t|s)$ is the conditional (on S=s) probability density of T. Note that $f_{T}(t)$ & $f_{T|S}(t|s)$ are not necessarily (usually not) the same, hence the use of subscripts (which will be suppressed when the context is obvious).

Note also that $R_T(t)$ does not equal $R_{T|S}(t|s)$.

$R_T(t)$ and $R_{T|S}(t|s)$ are related via

$$R_T(t) = \int_{t}^{\infty} \int_{\text{all } s \in S} f_{T|S}(t|s) f_S(s)dtds = \int_{\text{all } s \in S} R_{T|S}(t|s) f_S(s)ds$$

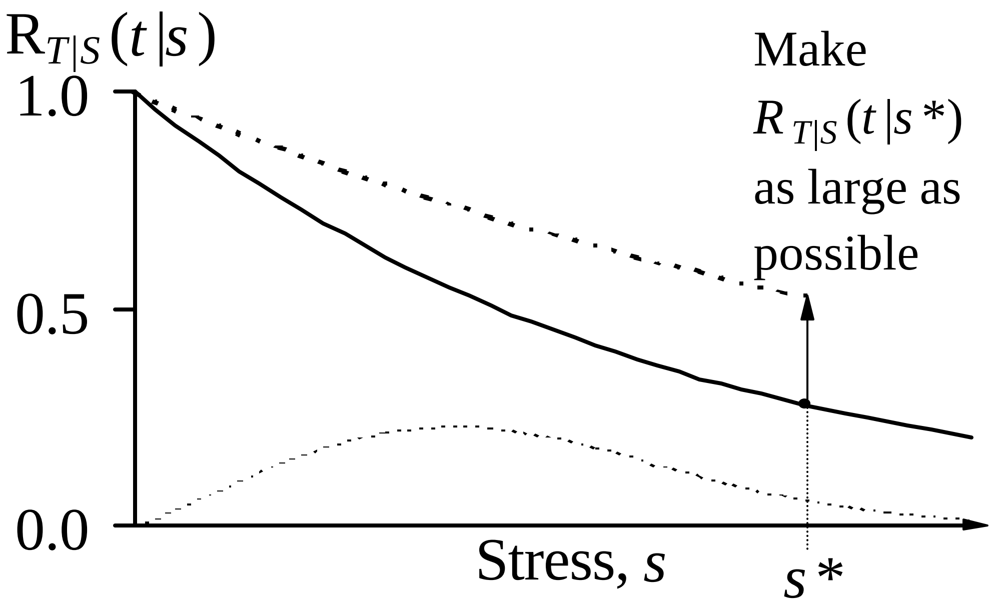
where $f_S(s)$ is the probability density function of the stress S in the field

Hence, $R_T(t)$ can be derived from $R_{T|S}(t|s)$, ***if*** we know $R_{T|S}(t|s)$ for *all* S-values, ***and*** if we know $f_S(s)$. This is usually not the case in automotive applications (as in medical studies).

Reliability Improvement

Fortunately, we can maximize $R_T(t)$ without needing to know it. In fact to maximize $R_T(.)$ we only have to maximize $R_{T|S}(.)$ across the range of S-values.

Furthermore, if $R_{T|S}(t|s)$ is a decreasing function of s for a fixed t, then to maximize $R_T(t)$ we just have to maximize $R_{T|S}(.)$ at a particular S-value (s^* say).



To demonstrate improvement in $R_{T|S}(t|s)$, it is usual to expose components to stresses on a rig or simulated stresses in CAE models, rather than use a vehicle.

Statistical engineering avoids over-engineering to achieve increase in $R_{T|S}(t|s)$.

To achieve reliability improvement using statistical engineering...

... use testing is a discovery mode, not just a verification one (experiments not just tests),

... anticipate the stresses that cause failure modes, and make the testing more traceable to real-world conditions,

... make the failure mode go away.

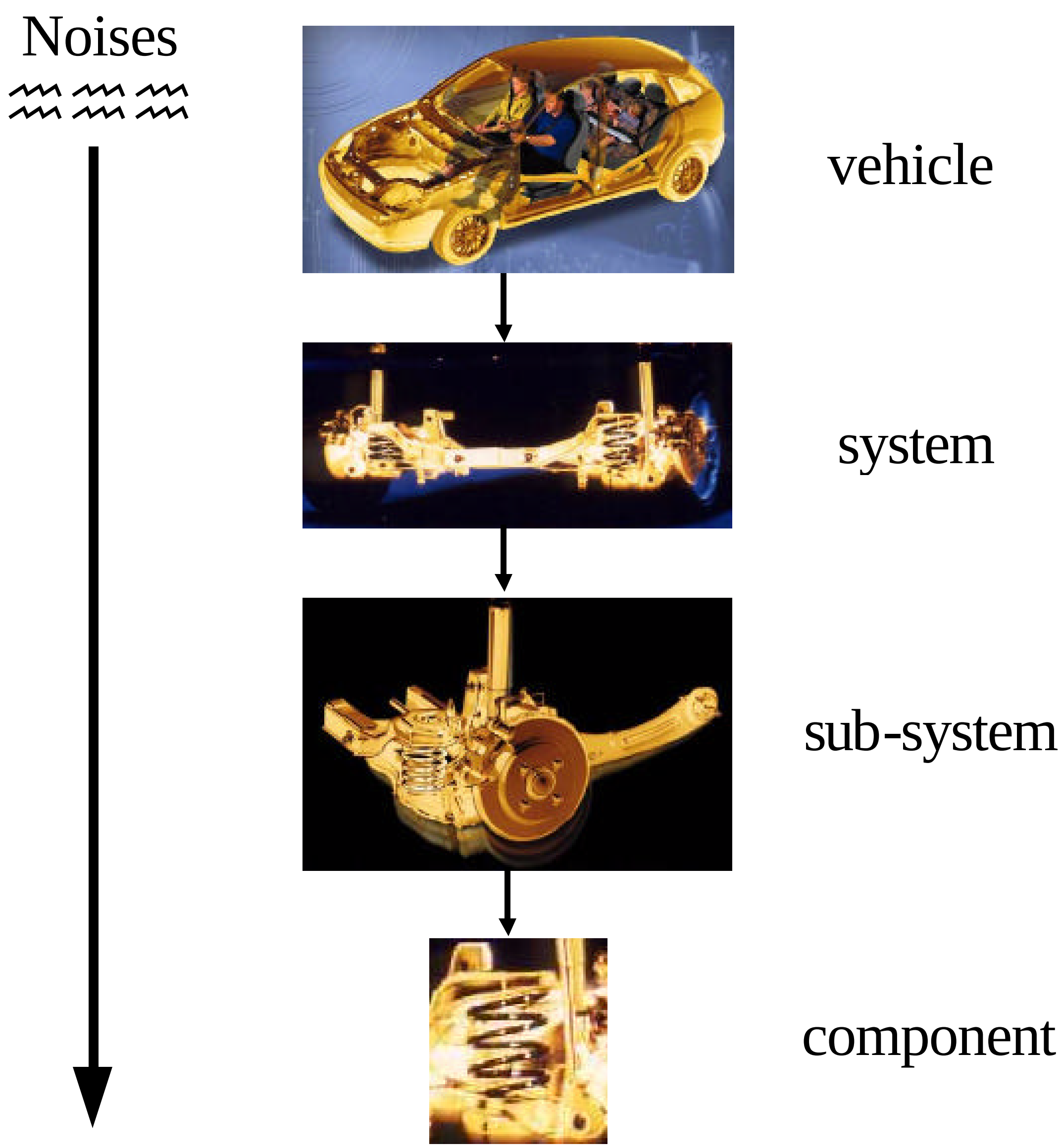
For reliability failures due to lack of robustness, the stresses can be subdivided in 5 categories. We call the stresses *Noise Factors*.

1. Piece-to-piece variation
2. Changes to component characteristics over time
3. Customer duty cycle
4. Climate & road conditions
5. Environment created by neighbouring components in the system.

Failure modes can be induced by many noise factors acting together, which adds another complication to the mathematical relationship between marginal & conditional reliability.

Ford takes as the “useful life period” 10 years or 150,000 miles. This is ***not*** a reliability target, but rather a period over which to dimension the noise factors in testing.

The idea is to cascade these noises down to the component level so that the reliability improvement work and reliability demonstration can take place there.



Reliability analysis using warranty data

Often in automotive applications, we have to rely on field data to make reliability assessments about components, and the field data itself is often only available through warranty data.

Some things to bear in mind about warranty data:-

Knowledge of failures is hampered because

- We have to assume that the claim date is equal to the failure date.
- We have to assume that a claim is actually a failure.
- Restrictions on warranty means that failures which happen *outside* warranty, are not known, and these are generally the high time/mileage failures.

Knowledge of the risk sets is hampered because

- Some vehicles are lost through accidents
- We do not know the mileage's of vehicles without a warranty claim.

Reliability analysis using Months to Failure as the life variable T is fairly straightforward because we know the sales dates of cars sold

Mileage is a little bit more tricky, because we don't know the mileage for unfailed cars.

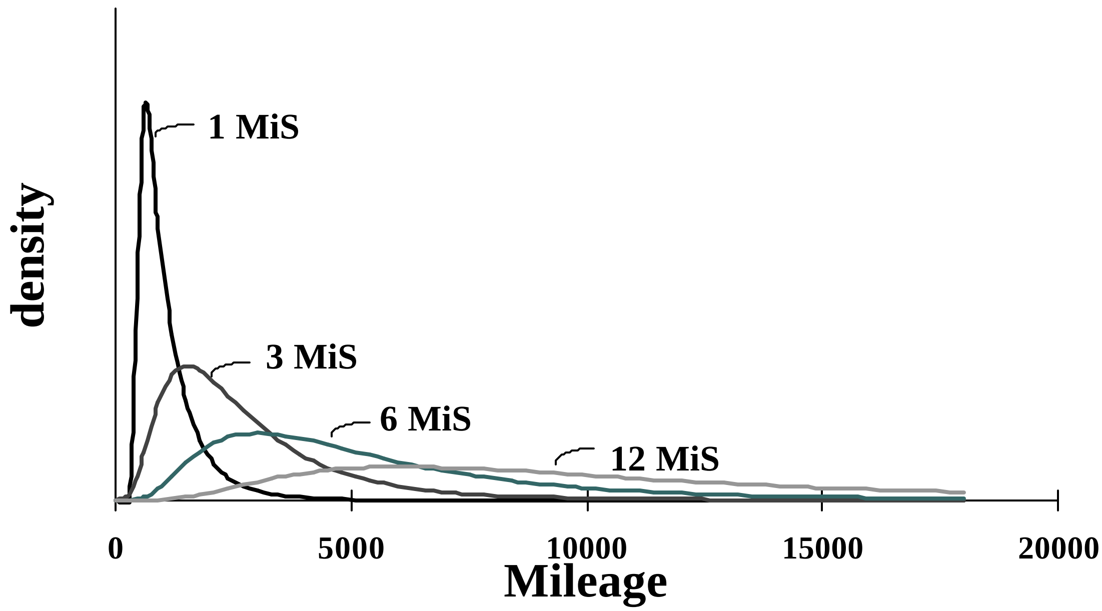
To a good approximation, the shape of the mileage (M) distribution is well described by the *lognormal* distribution. The average mileage accumulation (avg) *per month*, and the standard deviation of log(M) depends on the vehicle, and the market:-

Source: QAS	UK	D	US
Fiesta	550 mls	1350 km	-
Escort	750 mls	1450 km	750 mls
Mondeo	1050 mls	1850 km	-
Contour	-	-	950 mls
σ (log$M$)	0.60	0.55	0.65

$$\log(M|\text{service}=t) \sim \text{Norm}(\mu=\log(\text{avg}.t)-\tfrac{1}{2}\sigma^2, \sigma)$$

Using the lognormal distribution, the mileage distribution of cars without warranty claims can be inferred, and the risk sets adjusted accordingly.

Some lognormal distributions for 1,3,6,&12 Months in Service (MiS) :-

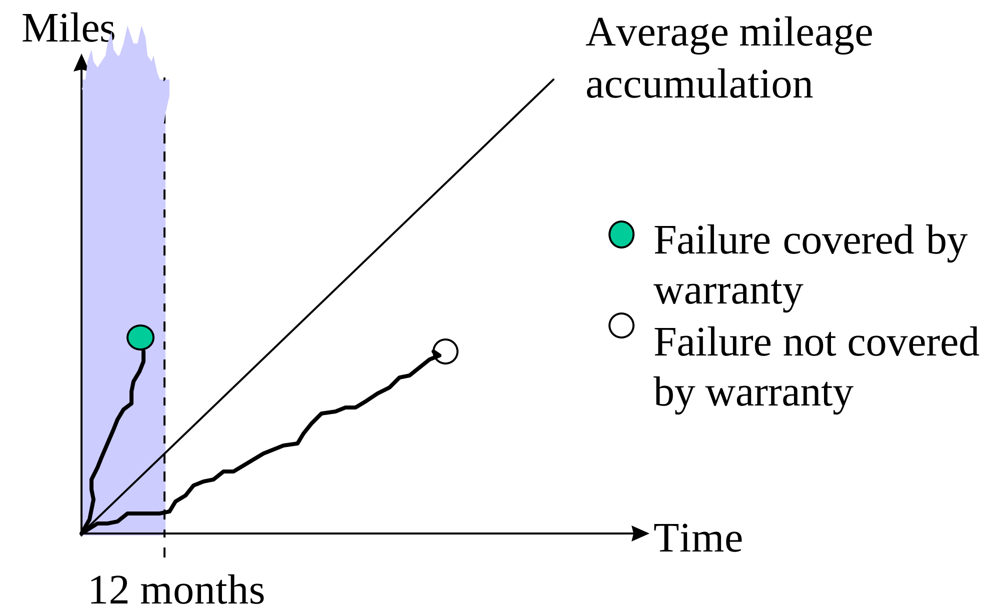


Working with Time or Mileage as the life variable leads to serious censoring problems (e.g. see Lawless, *et al*, *Lifetime Data Analysis*, Vol 1, 1995).

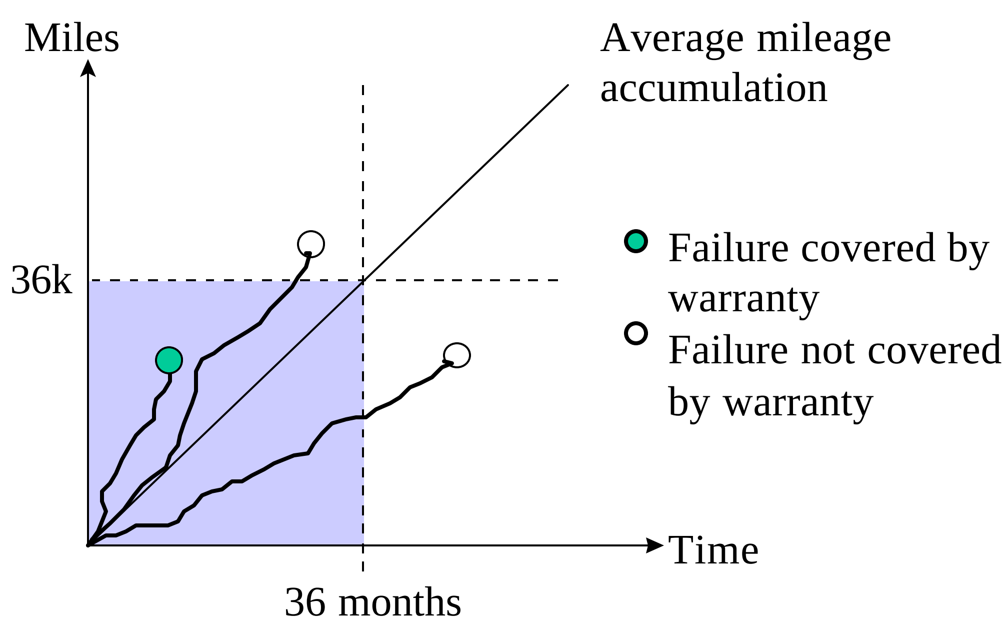
Maximum likelihood methods can be developed which overcome this censoring problem (Meeker & Davis, *work in progress*).

An illustration of the way warranty policies in EU and US affect failure information can be enlightening.

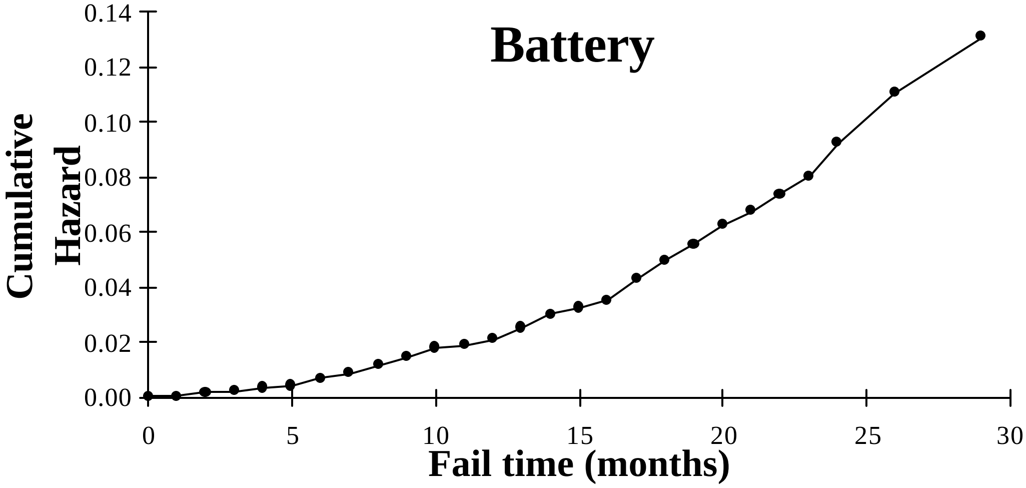
EU:



US:



Examples of Reliability Improvement



<u>Failure Mode:</u> Inability to deliver high rate discharge to 7.2v for 30s.

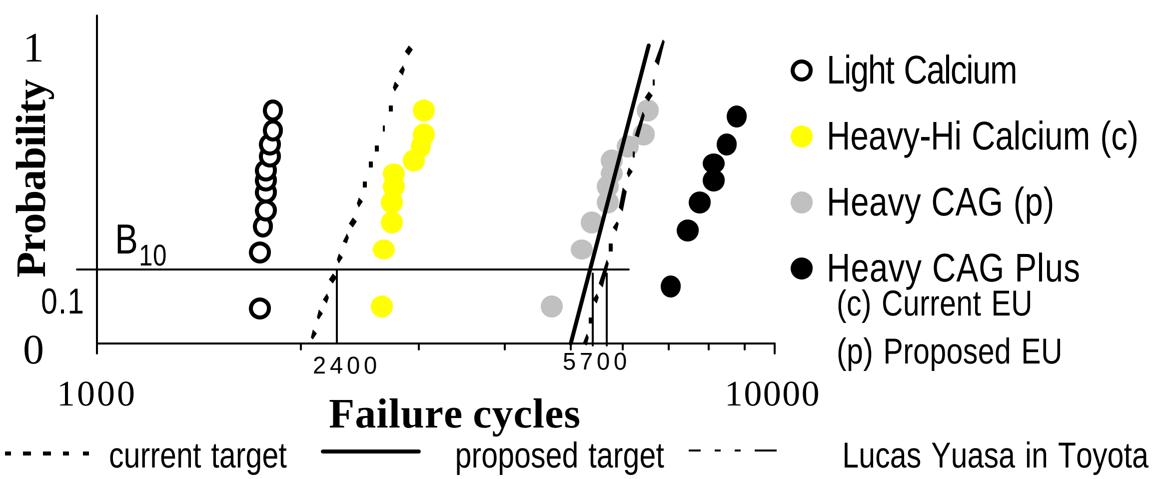
<u>N#1</u> randomly selected batteries to simulate material and dimensional variation (up to 9 per test)

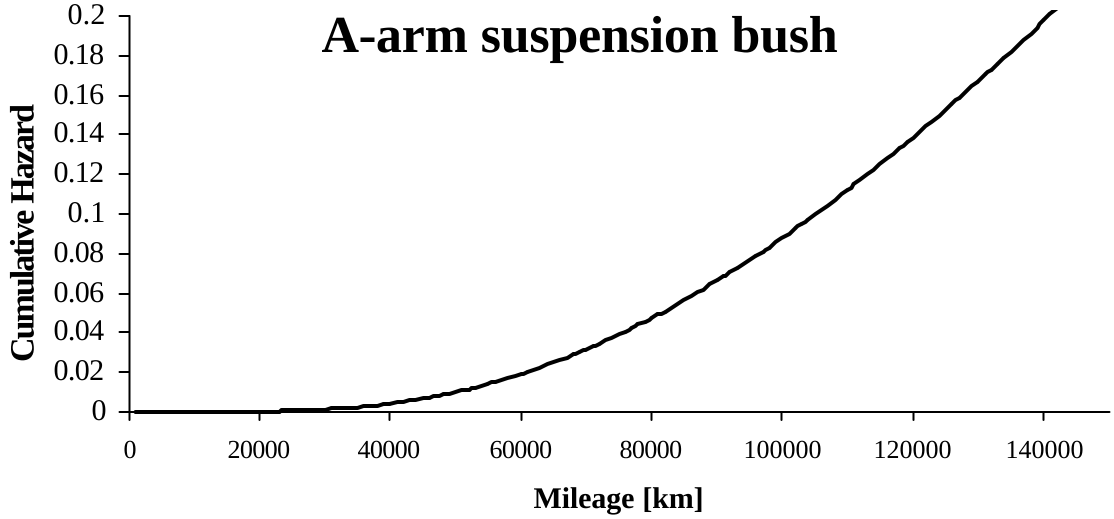
<u>N#2</u> generated by the test (material loss)

<u>N#3</u> deep charge (14.8v) and dis-charge (25A) cycles

<u>N#4</u> not applicable

<u>N#5</u> simulated: heat (75^{o}C) from the engine compartment, parasitic load.





<u>Failure Mode:</u> Rubber cracking and de-bonding

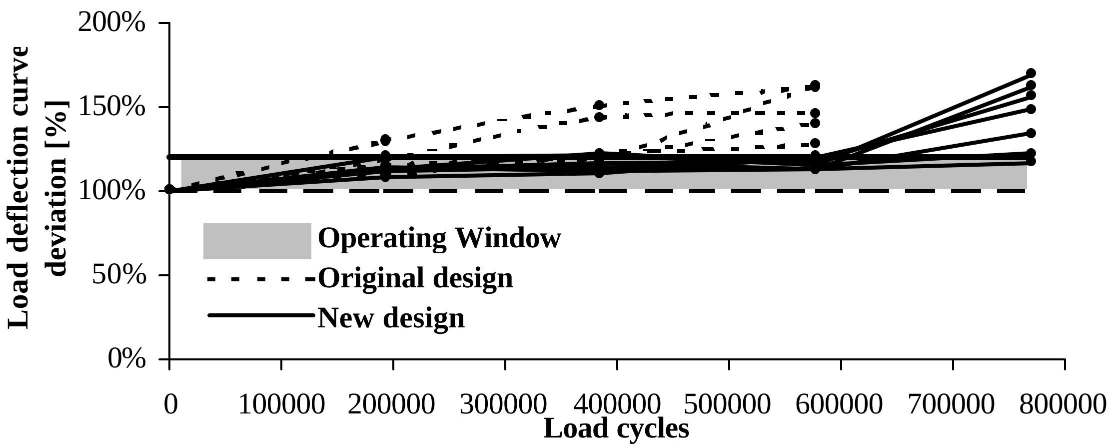
<u>N#1</u> captured by randomly selection

<u>N#2</u> is induced on new bushes directly by the test

<u>N#3</u> PASCAR loads applied (+4kn -> -2kN), cycled at accelerated frequency of 3Hz. Bush deflection angle $\pm 12^{o}$ at 1Hz.

<u>N#4</u> Cold start simulation (-25^{o}C), and corrosion chamber (Temp/Humidy/Salt)

<u>N#5</u> n/a.



Weibull regression models

(see Smith, *Reliability Engineering & System Safety*, Vol 34, 1991)

Consider a Weibull model for T where the scale parameter depends on a variable x, but the shape parameter does not, i.e.

$$f(t|x) = bq_x^{-b} t^{b-1} \exp\{-(t/q_x)^b\},$$

where q_x is a strictly positive function of x, e.g. $q_x = \exp(q_0 + q_1 x)$.

Use maximum likelihood to fit this Weibull regression model.

A simple, but very useful application is to demonstrate reliability improvement due to a design change. Using above notation

$x = 0$ for the prior (old level) design, and

$x = 1$ for the new (hopefully improved!) design.

Hence, the scale parameter q_x is equal to $\exp(q_0)$ for the old design, and $\exp(q_0+q_1x)$ for the new design, so the efficacy of the new design is measured through q_1, which we estimate from test data.

A more detailed application of Weibull regression

Consider data on failure times for 38 tyres tested on a rig. There are two covariates, x_1 and x_2 which relate to variation around two design nominals in production. There is no censoring.

Subject Index	Failure time	censoring indicator	*Covariates/Factors* x_0	x_1	x_2
1	234	1	1	-1.24	0.54
2	237	1	1	1.96	0.86
3	243	1	1	0.36	-0.01
4	248	1	1	0.46	1.11
5	258	1	1	4.26	0.57
6	304	1	1	0.66	0.01
7	320	1	1	4.06	-0.16
8	347	1	1	5.76	0.29
9	356	1	1	1.26	0.62
10	364	1	1	4.46	0.49
11	375	1	1	4.26	0.58
12	415	1	1	7.56	-0.43
13	425	1	1	-0.74	-0.47
14	436	1	1	1.96	0.07
15	452	1	1	5.56	-0.42
16	453	1	1	-5.74	0.21
17	465	1	1	-8.14	-0.36
18	467	1	1	-3.24	-0.21
19	477	1	1	-2.44	-0.09
20	497	1	1	-0.74	0.08
21	498	1	1	2.76	0.27
22	502	1	1	4.96	0.28
23	508	1	1	-5.24	-0.07
24	523	1	1	5.16	-0.24
25	556	1	1	9.36	-0.83
26	583	1	1	-5.54	-0.18
27	636	1	1	-7.74	-0.75
28	636	1	1	-4.54	0.51
29	646	1	1	-2.74	-0.34
30	666	1	1	7.26	-0.84
31	683	1	1	-9.74	0.02
32	703	1	1	-0.74	-0.79
33	711	1	1	0.76	-0.12
34	727	1	1	0.26	0.01
35	749	1	1	4.96	-0.66
36	760	1	1	-6.74	-0.38
37	834	1	1	-3.54	0.15
38	882	1	1	-9.24	0.50
39			1		

A natural model to fit initially is

$$\log(q_x) = q_0 + q_1 x_1 + q_2 x_2; \quad b_x = b$$

Note that the baseline reliability and hazard functions are given by a tyre with x_1=x_2=0, so that q_x=exp(q_0) in this case.

The fitted parameters are given in the following table:-

log-likelihood	Regression parameter	1/β	θ_0	θ_1	θ_2
-7.293	**Fitted Coefficient**	0.259	6.300	-0.036	-0.373
	Standard Error	0.0322	0.0444	0.0090	0.1058
	t-ratio	8.02	141.74	-3.94	-3.53
	Significant?	#N/A	#N/A	yes	yes

Underneath each coefficient is its *standard error*. The *t-ratio* is the ratio of the fitted coefficient to its standard error; if the ratio is bigger than ~2, the coefficient is deemed significant. In this example both covariates have significant (negative) effects on the q_x, so a lower value is better.

The baseline Weibull distribution has the following parameters:-

parameter	value	95% confidence interval	
q	544.7	(498.3	595.3)
b	3.87	(3.10	5.15)
B1 Life	165.8	(118.39	232.07)
B5 Life	252.7	(199.50	319.99)
B10 Life	304.4	(250.80	369.36)

Note that *b*>1, so we have a failure mechanism with IFR.

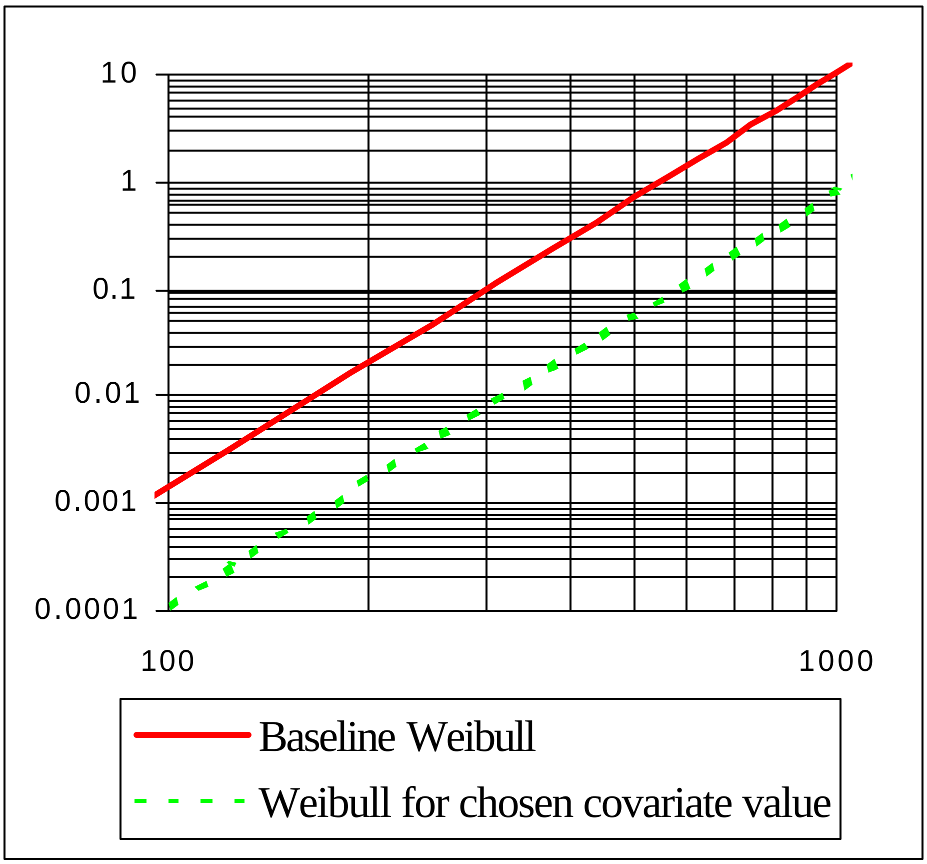
The fitted model can be used to make predictions, e.g.

	Covariates/Factors		
	x_0	x_1	x_2
Covariate Values	1	-8	-1
Memo	q_0	q_1	q_2
Fitted Coefficients	6.300	-0.036	-0.373
Significant?	#N/A	yes	yes

For an x_1 value of -8, and an x_2 value of -1, the fitted Weibull parameters are

parameter	value	95% confidence interval	
q	1051.6	(763.7	1448.0)
b	3.87	(3.10	5.15)
B1 Life	320.0	(201.7	507.8)
B5 Life	487.8	(329.6	722.0)
B10 Life	587.6	(407.2	848.0)

A probability plot of the baseline Weibull compared to the Weibull with x_1=-8 and x_2=-1 looks as follows:-



So that reliability improvement can be demonstrated by the Weibull plot moving to the right.

Design of experiments in reliability applications

The previous example on tyres illustrated an analysis based on *passive observation* of the tested components. However, intervention via *directed experimentation* also has a major role to play in reliability improvement.
Two main technical problems arise in the reliability area when standard methods of data analysis are employed.

1. Failure time data are rarely normally distributed, so standard analysis tools which rely on symmetry, e.g. Normal plots, don't work too well.
2. Censoring.

The first problem can be overcome by considering a *transformation* of the fail times to make them approximately normal - the log transformation is usually a good choice. The exact form of the fail time distribution is not important because we are looking for effects that *improve* reliability, rather than exact predictions of the reliability itself.

The second problem of censoring is a little bit more tricky, but can be dealt with by iteration as follows:-

1. Choose a basic model to fit to the data.
2. Fit the model to the data, treating the censor times as failure times.
3. Using this model, make a *conditional* prediction for the unobserved fail times for each censored observation. Conditional because the actual failure time must be consistent with the censoring mechanism.
4. Replace censor times with the fail time predictions from step 3.
5. Go back to Step 2.

Eventually this process will converge i.e. the predictions for the fail times of the censorings will stop changing from one iteration to the next.
If necessary the process can be tried with several model choices for step 1.
In fact, the algorithm described above leads to the same results as maximum likelihood estimation.

(Hamada & Wu, *Technometrics*, Vol 33,1991)

Example on Transmission Centre Shafts

Consider a designed experiment to improve the reliability of the centre shaft of an automatic transmission.

7 design and manufacturing factors were considered (levels 1/2/3 in brackets)

A. Profile of spline end (spherical/grooved/-)
B. Amount of annealing (nil/1 hr/4 hrs)
C. Shaft diameter (16.1mm/17.7mm/18.8mm)
D. Intensity of shot (3 Almen/6 Almen/9 Almen)
E. Coverage of shot (200%/400%/600%)
F. Tempering Temperature (140^o/160^o/180^o)
G. Shot Blasting (with/without/-)

These factors were allocated to an orthogonal array design, and two transmission centre shafts were made for each specified configuration. The shafts were rig tested until failure, or 2×10^6 cycles, whichever came first. 17 transmission shafts were censored in all at the 2×10^6 cycles bogey (indicating that the test could probably have been accelerated more).

Because of the large proportion of censoring, the algorithm took a long time to converge (about 14 iterations).

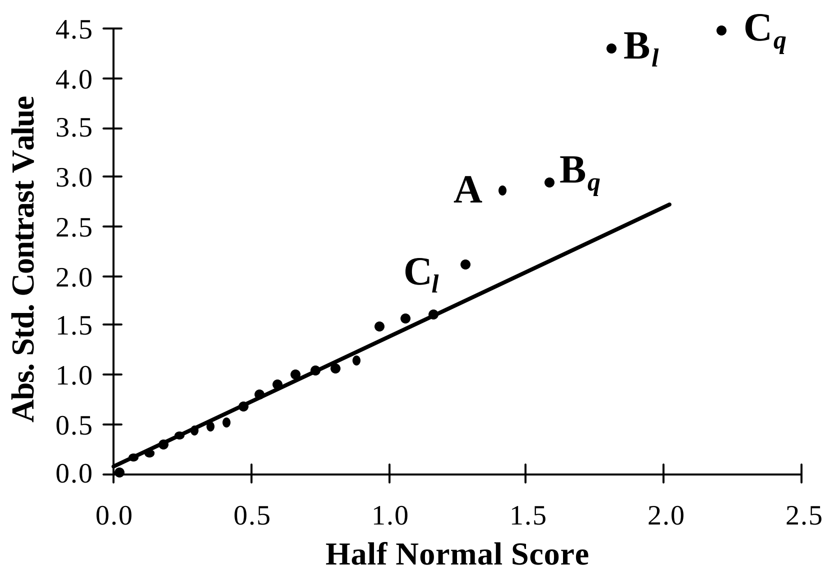
The “data” before and after convergence after fitting a main effects model is as follows:-

Config	Factors A	B	C	D	E	F	G	Original Data		"Data" after 14 iterations	
1	1	1	2	2	2	2	1	322	2000+	322	2794
2	1	2	2	1	1	1	2	95	95.4	95	95.4
3	1	3	2	3	3	3	1	2000+	125	3355	125
4	1	1	1	2	1	3	1	747	414	747	414
5	1	2	1	1	3	2	1	821	192	821	192
6	1	3	1	3	2	1	2	2000+	2000+	4145	4145
7	1	1	3	1	2	1	1	972	2000+	972	4785
8	1	2	3	3	1	3	1	2000+	1920	4141	1920
9	1	3	3	2	3	2	2	2000+	2000+	12196	12196
10	2	1	2	3	3	1	1	739	285	739	285
11	2	2	2	2	2	3	2	1080	634	1080	634
12	2	3	2	1	1	2	1	2000+	1940	6355	1940
13	2	1	1	1	3	3	2	2000+	1790	4229	1790
14	2	2	1	3	2	2	1	2000+	617	4278	617
15	2	3	1	2	1	1	1	2000+	2000+	18671	18671
16	2	1	3	3	1	2	2	1380	1110	1380	1110
17	2	2	3	2	3	1	1	2000+	2000+	4786	4786
18	2	3	3	1	2	3	1	2000+	2000+	30964	30964

Censored observations are marked with a +.

The data after convergence can now be used with standard data analysis tools.

For example, a half-normal plot of the effects from the experiment yields the following:-



with effects plotting above the line indicating these factors have a significant effect on lifetime.

(more details in Davis, *Lifetime Data Analysis* Vol. 1 1995).

Reliability improvement through parameter design

The achievement of higher reliability can also be viewed as an improvement to Robustness. Robustness is defined as reduced sensitivity to the *noise factors*, which can sometimes be achieved by changing nominals of design variables (*control factors*).

Reminder of the 5 main categories:-

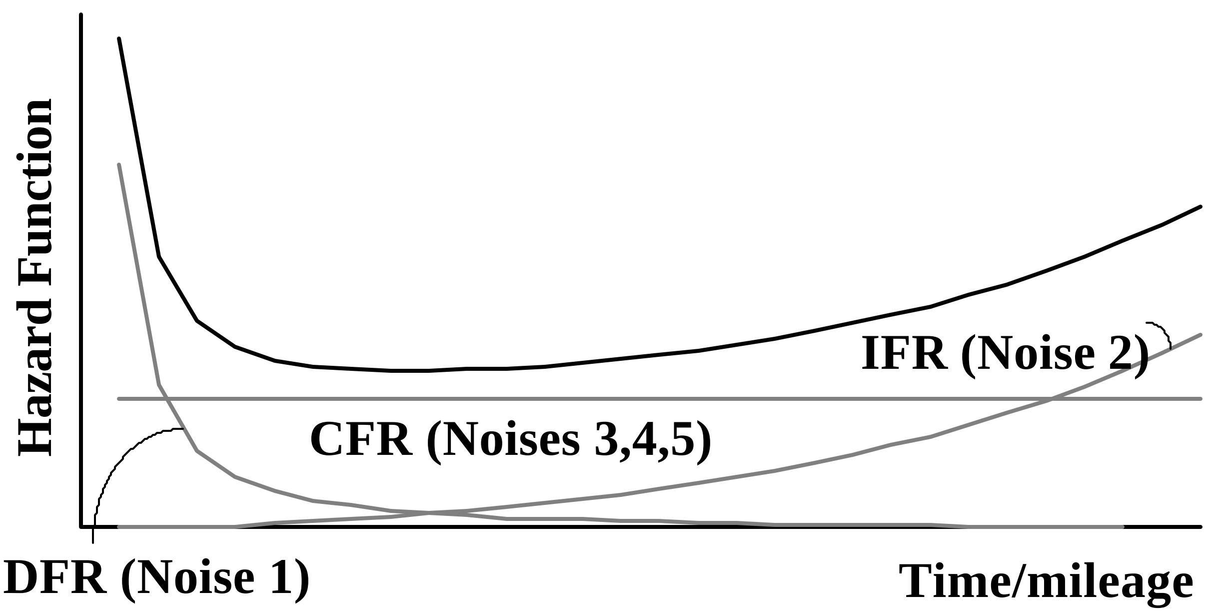
1. Piece-to-piece variation
2. Changes to component characteristics over time
3. Customer duty cycle
4. Climate & Road conditions
5. Environment created by neighbouring components in the system.

Typically, noises in categories 3,4,&5 can induce noises in category 2. If the function of the component can be made robust to noises in category 2, then the component will, by definition, be more reliable.

Bathtub Curve of the hazard function

TPD's experience:- noise category 1 contributes to DFR failure rates (infant mortallity), category 2 to IFR, and categories 3, 4, & 5 to CFR or "useful life" problems.

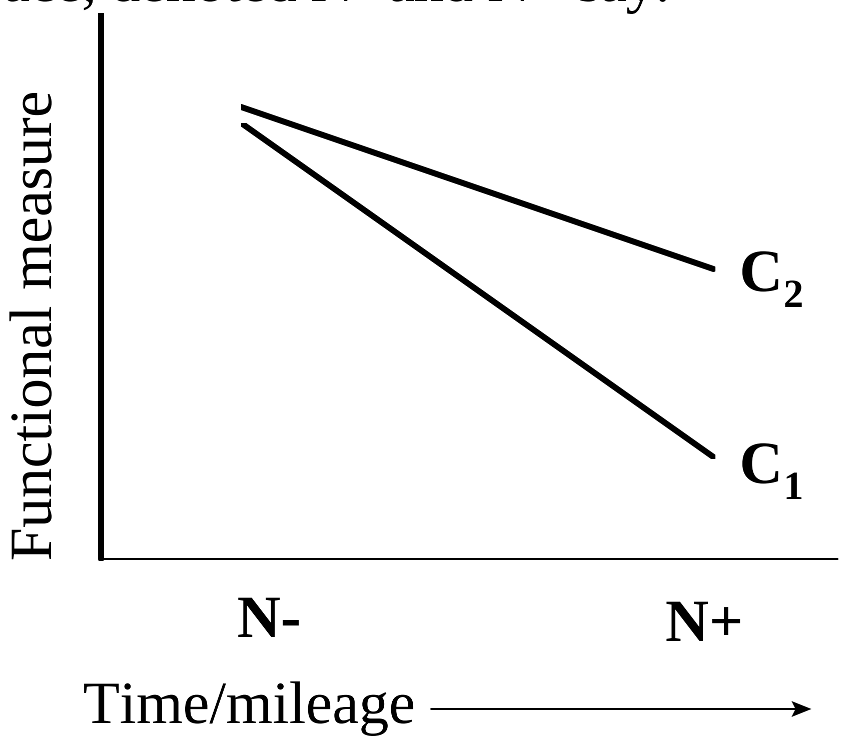
Thinking about the noise factors in this way produces the well known bathtub curve for the hazard function



Often, knowing the type of failure rate that is acting on our component can give a clue as to the offending noise factor, and hence lead to a root cause analysis of the failure mechanism.

Components can be made robust to noises by experimenting with *control factors*.

The idea (as in robustness generally) is to look for interactions between control and noise factors. The reliability connection is made if there is a life variable interpretation between the extremes of the noise space, denoted N- and N+ say.



(e.g. N- might be a “new” condition, and N+ “old)

Note that the functional measure is *not* failure time, but some ideal function of the system. C_1 and C_2 represent two settings of a control factor. A design with C_2 is more robust to noise than one with C_1 and is therefore more reliable.

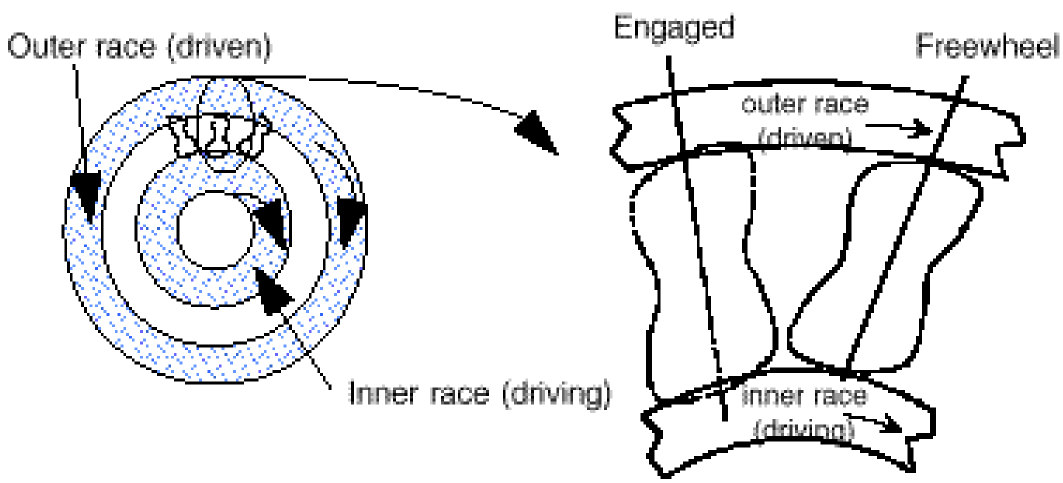
A parameter design layout in reliability applications follows the pattern for parameter design studies generally, for example:-

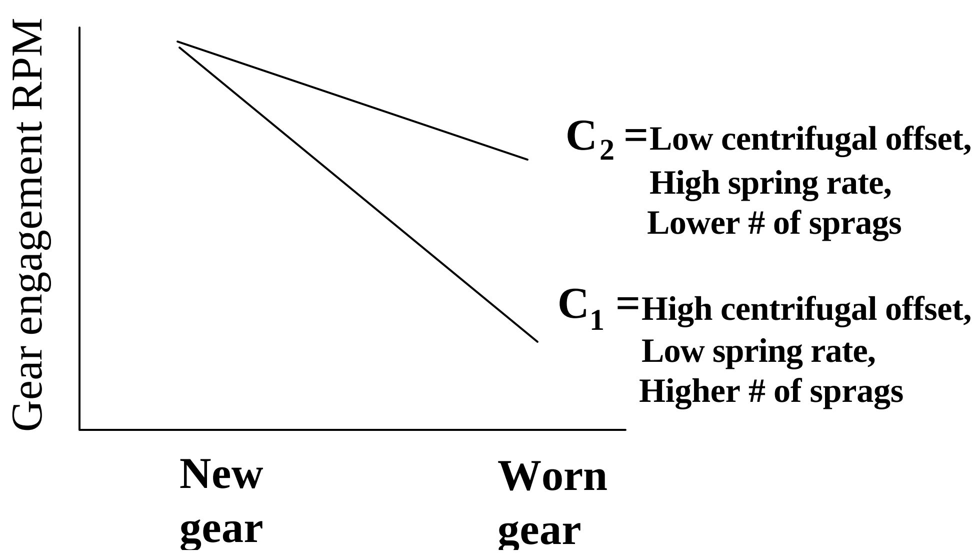
	Control Factors					Noise Factors	
						time/mileage →	
Config-uration	A	B	C	...	G	N- (new)	N+ (old)
1	-	-	-		-	y_{1-}	y_{1+}
2	+	-	-		+	y_{2-}	y_{2+}
3	-	+	-		+	y_{3-}	y_{3+}
4	+	+	-		-	y_{4-}	y_{4+}
5	-	-	+		+	y_{5-}	y_{5+}
6	+	-	+		-	y_{6-}	y_{6+}
7	-	+	+		-	y_{7-}	y_{7+}
8	+	+	+		+	y_{8-}	y_{8+}

The idea of experimental layouts of this type is to look for interactions between Control Factors and the Noise Factors, which lead to configurations with minimum difference between the *y*-values.

Example

The sprag one-way clutch is a sub-assembly in ATX's to allow non-synchronous shifting. It is designed to transmit torque in direction of relative motion, but not in the other direction, thereby allowing slippage/free wheeling if desired.





What we are doing in

- training engineers to be “statistical engineers”
- integrating reliability plans into our engineering process.
- working on long standing reliability problems.
- encouraging our suppliers to adopt some of these ideas.
- continuous management reviews