Solutions 3BA1 2005 summer

		c	not(c)	
	p	0.1	0.08	
	Np	0.25	0.06	L
	p	0.32	0.07	not L
1	NP	0.08	0.04	

P(p) 0.57
P(P|C)= 0.56 P(C|P) 0.736842

Not independent as P(P)<>P(P|C)

P(C|LandP) 0.555555556
Know=(Land shape and a Person) prob that its P(colour)

P(LuCbar|P) 0.438596

P(P and Cbar)=0.15 0.15

P(at least 1) =1- P(none) =1-0 1-(1-0.15)^3 0.385875

P(L) 0.49
P(Lbar) 0.51
p
0.5002

Qyestion 2
(a) P(S|G)=0.02, P((S|Gbar)=0.9

(b) Binomial model
P(X<=2)= BINOMIST(2,10,0.4,1)

P(Gbar|S)=(p(S|Gbar)*P(Gbar))/P(S)
P(S)=P(S|G)P(G)+P(S|gbar)P(gbar)

P(S)= 0.02*0.6+0.9*0.4
0.372
0.967742

Two Gs reported S bar 0.9604
Gbar reported as S bar 0.1
0.09604

P(Gbar|S)=0.02
0.936524

(a)
i life 0.3
P(T>3) 0.40657
exp(-0.9)

ii It means that they don't age
P(last another T|age A) is independent of A

iii Thus exp(-0.3) 0.740818

(b)

i Gaps in PP are exponentially distribtuted.
Time to kth event = sum of k indep. Exponentials
Assumption - image sizes are independent.

ii mean here is 6.666667

So Poisson(4,6.6667,1)

1-poisson(9,6.6667,1)

iii Solve

R= F(X)

X=4* (ln(1-R))^2

R	W	1+W
0.753	7.82172	8.82172
0.147	0.101119	1.101119

4

(a) The standard error = 11.25

tstat 0.888889

As <1 not significant no need to look up

(b) The test in (a) tests only that the average RPM of motors is 7000
Even if true this gives no guarantee of usability.
eg. Half the motors could have rpm=6000 and half 8000.

(c) phat 0.8

se phat 0.023094

So CI 0.753812 0.846188

(d) phat2 0.888889

Diff p 0.088889

pooled P 0.833333

se(diff) 0.086066

			pvalue
Zratio	1.032796	Not sig	0.15085

Q5

(a) To analyze these data require that the quality measurments be independent.

(b) quality 0.234 +0.0009111*n n=generations

at n=2000

2.0559 which is nonsense

Quality cannot be greater than 1

2000 is an extrapolation beyond the range of the data the LINEAR not hold that far out.

The scatter plot shows and the residual plot strangly confrims CURVATURE.

(c) R-squared is superior
The plot look more like they should
The scatter plot look like a straight line.

	0.007828	0.000484
CI	0.0068604	0.008796

at n=1000		logit
-2.15844	0.007828	5.66956

ssx	941000		
((x-xbar)^2	324900	1/n	0.1

s	mult	s.E.
0.4693	1.730975	0.812347

Low Logit	Hi logit
4.044866505	7.294253

Low qual	Hi qual	qual=1/(1+exp(-logit))
0.982789351	0.999321	

Quality = 1/(1+exp(L(x))

Q6 coeficients= estimates of constant and slope
(a) SD(estimates)
T-value - t-stat for testing coeficient =0
Constant not significant
Size is significant.
p-value convertions of t-stats into a probability.
Sig if p<0.05

prediction

	869.8405	secs

ssx	10200
(Xbar-60)^2	100

SE	58.49755

752.8454054	986.8356

(b) (i)
Ouput 4 shows that neither the SITE or slope variable are significant

However Output 3 shows that the sitslope vartiable is significant if site is omitted

In output 2 The site variable is nearlyy significant

SIZE is significant in all outputs.

It therefore looks like output 3 give the correct picture
No differences in the constant from site to site but the slope - = secs/MB is different

the equation looks like

104.3	+ 12.8 secs/MB	if site 1
104.3	+ 8.8 secs/MB	if site 2

(ii) F-test

		df
SSE(4)	1.13	21
SSE(1)	1.4246	23

Numerator	0.1473
Denom	0.05381
Fstat	2.737434

Critical value F(2,21)

3.4668 Not signif

F test is used to ascertain whether several coeficients can be simulatniuosly regarded as 0.