Risk Analysis of the Space Shuttle: Pre-Challenger Prediction of Failure

Marie-Gabrielle Dondon

11/10/2018

This is a SAS version based on the Rmd document of Arnaud Legrand available at https://app-learninglab.inria.fr (https://applearninglab.inria.fr/gitlab/moocrr-session1/moocrr-reproducibility-study/blob/5e298cd5b653bcb31b6f9fedcb7f75e800c2dc98/src/R/challenger.Rmd) and the SAS program of Michael Friendly available at https://dokumen.tips (https://dokumen.tips/documents/categorical-data-analysis-with-graphics.html) (cf. page 52).

In this document we reperform some of the analysis provided in *Risk Analysis of the Space Shuttle: Pre-Challenger Prediction of Failure* by *Siddhartha R. Dalal, Edward B. Fowlkes, Bruce Hoadley* published in *Journal of the American Statistical Association*, Vol. 84, No. 408 (Dec., 1989), pp. 945-957 and available at http://www.jstor.org/stable/2290069 (http://www.jstor.org/stable/2290069).

On the fourth page of this article, they indicate that the maximum likelihood estimates of the logistic regression using only temperature are: $\hat{\alpha}$ = **5.085** and $\hat{\beta}$ = **-0.1156** and their asymptotic standard errors are $s_{\hat{\alpha}}$ = **3.052** and $s_{\hat{\beta}}$ = **0.047**. The Goodness of fit indicated for this model was G^2 = **18.086** with **21** degrees of freedom. Our goal is to reproduce the computation behind these values.

The DATA step below reads the data on the number of O-ring failures and temperature for the 23 flights forwhich information was available before the Challenger launch. Our interest here is in predicting the likelihood of failures at low temperatures.

In [1]:
```
ods title;
ods noproctitle;
title "NASA Space Shuttle O-Ring Failures";

data nasa;
  input Flight $ Date :mmddyy8. Count Temperature Pressure Malfunc
tion;
  format Date mmddyy8.;
  cards;
1    04/12/81  6  66   50  0
2    11/12/81  6  70   50  1
3    03/22/82  6  69   50  0
5    11/11/82  6  68   50  0
6    04/04/83  6  67   50  0
7    06/18/82  6  72   50  0
8    08/30/83  6  73  100  0
9    11/28/83  6  70  100  0
41B  02/03/84  6  57  200  1
41C  04/06/84  6  63  200  1
41D  08/30/84  6  70  200  1
41G  10/05/84  6  78  200  0
51A  11/08/84  6  67  200  0
51C  01/24/85  6  53  200  2
51D  04/12/85  6  67  200  0
51B  04/29/85  6  75  200  0
51G  06/17/85  6  70  200  0
51F  07/29/85  6  81  200  0
51I  08/27/85  6  76  200  0
51J  10/03/85  6  79  200  0
61A  10/30/85  6  75  200  2
61B  11/26/85  6  76  200  0
61C  01/12/86  6  58  200  1
;
run;

data nasa;
  set nasa;
  label Malfunction = "Number of O-ring failures"
        Temperature = "Temperature (deg F)";
run;

proc print data=nasa (obs=5);
  id Flight;
run;
```

```
SAS Connection established. Subprocess id is 3488
```

Out[1]:

NASA Space Shuttle O-Ring Failures

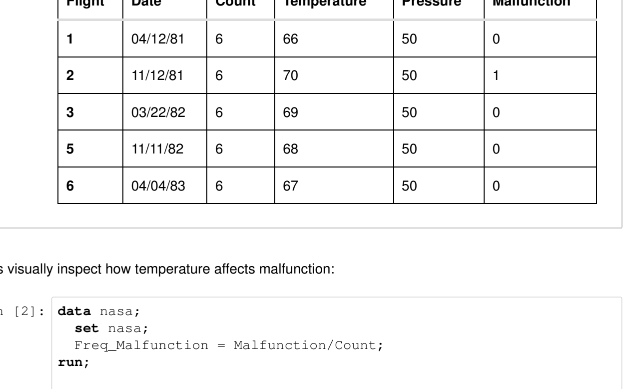
Flight	Date	Count	Temperature	Pressure	Malfunction
1	04/12/81	6	66	50	0
2	11/12/81	6	70	50	1
3	03/22/82	6	69	50	0
5	11/11/82	6	68	50	0
6	04/04/83	6	67	50	0

Let's visually inspect how temperature affects malfunction:

In [2]:
```
data nasa;
  set nasa;
  Freq_Malfunction = Malfunction/Count;
run;

proc gplot data=nasa;
  plot Freq_Malfunction * Temperature / haxis=axis1 hminor=0
                                        vaxis=axis2 vminor=0;
  symbol1 v=dot i=none c=blue h=2;
  axis1 order=(50 to 85 by 5);
  axis2 order=(0 to 1 by 0.2) label=(angle=90 'Estimated Failure p
robability');
run;
quit;
```

Out[2]:



To obtain predicted probabilities for observations not in the original sample, create an additional data set which contains values for the independent variables in the extrapolation sample, and join these observations to the actual data set. The response variable (Malfunction) will be missing for the extrapolation sample.

Obtain predicted values for 30-80 degrees

In [3]:
```
data temp;
  input Temperature @@;
  cards;
31 30 35 40 45 50 55 60 65 70 75 80
;
run;

data nasa2;
  set nasa temp;
run;

proc print data=nasa2;
  id Flight;
run;
```

Out[3]:

NASA Space Shuttle O-Ring Failures

Flight	Date	Count	Temperature	Pressure	Malfunction	Freq_Malfun
1	04/12/81	6	66	50	0	0.00000
2	11/12/81	6	70	50	1	0.16667
3	03/22/82	6	69	50	0	0.00000
5	11/11/82	6	68	50	0	0.00000
6	04/04/83	6	67	50	0	0.00000
7	06/18/82	6	72	50	0	0.00000
8	08/30/83	6	73	100	0	0.00000
9	11/28/83	6	70	100	0	0.00000
41B	02/03/84	6	57	200	1	0.16667
41C	04/06/84	6	63	200	1	0.16667
41D	08/30/84	6	70	200	1	0.16667
41G	10/05/84	6	78	200	0	0.00000
51A	11/08/84	6	67	200	0	0.00000
51C	01/24/85	6	53	200	2	0.33333
51D	04/12/85	6	67	200	0	0.00000
51B	04/29/85	6	75	200	0	0.00000
51G	06/17/85	6	70	200	0	0.00000
51F	07/29/85	6	81	200	0	0.00000
51I	08/27/85	6	76	200	0	0.00000
51J	10/03/85	6	79	200	0	0.00000
61A	10/30/85	6	75	200	2	0.33333
61B	11/26/85	6	76	200	0	0.00000
61C	01/12/86	6	58	200	1	0.16667
	.	.	31	.	.	.
	.	.	30	.	.	.
	.	.	35	.	.	.
	.	.	40	.	.	.
	.	.	45	.	.	.
	.	.	50	.	.	.
	.	.	55	.	.	.
	.	.	60	.	.	.
	.	.	65	.	.	.
	.	.	70	.	.	.
	.	.	75	.	.	.
	.	.	80	.	.	.

In the PROC LOGISTIC step, we use the events/trials syntax to indicate the number of failures and number of trials. The observations in the extrapolation sample are not used in fitting the model, yet the procedure produces predicted probabilities and logits (as long as the independent variable(s) are non-missing).

In [4]:
```
ods output OddsRatios=OddsRatios ParameterEstimates=ParameterEstim
ates GoodnessOfFit=GoodnessOfFit;
proc logistic data=nasa2 nosimple;
  model Malfunction/Count = Temperature / scale=none;
  output out=results p=predict l=lower u=upper;
run;
```

Out[4]:

NASA Space Shuttle O-Ring Failures

Informations sur le modèle		
Table	WORK.NASA2	
Variable de réponse (Evénements)	Malfunction	Number of O-ring failures
Variable de réponse (Expériences)	Count	
Modèle	logit binaire	
Technique d'optimisation	Score de Fisher	

Nb d'observations lues	35
Nb d'observations utilisées	23
Somme des fréquences lues	138
Somme des fréquences utilis	138

Profil de réponse		
Valeur ordonnée	**Résultat binaire**	**Fréquence totale**
1	Evénemen	9
2	Non-évén	129

Note: 12 observations were deleted due to missing values for the response or explanatory variables.

Etat de convergence du modèle
Critère de convergence (GCONV=1E-8) respecté.

Statistique d'adéquation de la déviance et de Pearson				
Critère	**Valeur**	**DDL**	**Valeur/DDL**	**Pr > Khi-2**
Ecart	18.0863	21	0.8613	0.6435
Pearson	29.9803	21	1.4276	0.0924

Nombre d'observations d'événements/expériences : 23

Statistiques d'ajustement du modèle			
		Constante et covariables	
Critère	**Constante uniquement**	**Log-vraisemblance**	**Log-vraisemblance complète**
AIC	68.540	64.396	35.647
SC	71.468	70.251	41.501
-2 Log	66.540	60.396	31.647

Test de l'hypothèse nulle globale : BETA=0			
Test	**Khi-2**	**DDL**	**Pr > Khi-2**
Rapport de vrais	6.1440	1	0.0132
Score	6.7696	1	0.0093
Wald	6.0435	1	0.0140

Estimations par l'analyse du maximum de vraisemblance					
Paramètre	**DDL**	**Estimation**	**Erreur type**	**Khi-2 de Wald**	**Pr > Khi-2**
Intercept	1	5.0850	3.0525	2.7751	0.0957
Temperature	1	-0.1156	0.0470	6.0435	0.0140

Estimations des rapports de cotes			
Effet	**Valeur estimée du point**	**95% Intervalle de confiance de Wald**	
Temperature	0.891	0.812	0.977

Association des probabilités prédites et des réponses observées			
Pourcentage concordant	65.4	**D de Somers**	0.382
Pourcentage discordant	27.1	**Gamma**	0.413
Pourcentage lié	7.5	**Tau-a**	0.047
Paires	1161	**c**	0.691

The printed output indicates that the 12 new observations were not used in the analysis.

In [5]:
```
proc print data=OddsRatios label;
  id Effect;
  title "Odds Ratios";
run;
```

Out[5]:

Odds Ratios

Effet	Estimation du rapport de cotes	Borne inférieure de l'IC à 95% pour le rapport de cotes	Borne supérieure de l'IC à 95% pour le rapport de cotes
Temperature	0.891	0.812	0.977

The odds ratio, 0.891, is interpreted to mean that each increase of 1° in temperature decreases the odds of a failure by 11%!

In [6]:
```
proc print data=ParameterEstimates label;
  id variable;
  format estimate stderr 8.5;
  var _numeric_;
  title "Parameter Estimates";
run;
```

Out[6]:

Parameter Estimates

Variable	DDL	Estimation	Erreur type	Khi-2 de Wald	Pr > Khi-2
Intercept	1	5.08498	3.05249	2.7751	0.0957
Temperature	1	-0.11560	0.04702	6.0435	0.0140

The maximum likelyhood estimator of the intercept and of Temperature are thus $\hat{\alpha}$ = **5.0850** and $\hat{\beta}$ = **-0.1156** and their standard errors are $s_{\hat{\alpha}}$ = **3.052** and $s_{\hat{\beta}}$ = **0.04702**.

In [7]:
```
proc print data=GoodnessOfFit label;
  id criterion;
  title "Goodness Of Fit";
run;
```

Out[7]:

Goodness Of Fit

Critère	DDL	Khi-2	Khi-2/DDL	Pr > Khi-2
Ecart	21	18.0863	0.8613	0.6435
Pearson	21	29.9803	1.4276	0.0924

The Residual deviance corresponds to the Goodness of fit G^2 = **18.086** with **21** degrees of freedom.

The output data set results contains the predicted probability of a failure at each temperature and upper and lower confidence 95% limits for this probability.

In [8]:
```
proc print data=results;
  id Flight;
  title "Predicted probabilities of a failure";
run;
```

Out[8]:

Predicted probabilities of a failure

Flight	Date	Count	Temperature	Pressure	Malfunction	Freq_Malfun
1	04/12/81	6	66	50	0	0.00000
2	11/12/81	6	70	50	1	0.16667
3	03/22/82	6	69	50	0	0.00000
5	11/11/82	6	68	50	0	0.00000
6	04/04/83	6	67	50	0	0.00000
7	06/18/82	6	72	50	0	0.00000
8	08/30/83	6	73	100	0	0.00000
9	11/28/83	6	70	100	0	0.00000
41B	02/03/84	6	57	200	1	0.16667
41C	04/06/84	6	63	200	1	0.16667
41D	08/30/84	6	70	200	1	0.16667
41G	10/05/84	6	78	200	0	0.00000
51A	11/08/84	6	67	200	0	0.00000
51C	01/24/85	6	53	200	2	0.33333
51D	04/12/85	6	67	200	0	0.00000
51B	04/29/85	6	75	200	0	0.00000
51G	06/17/85	6	70	200	0	0.00000
51F	07/29/85	6	81	200	0	0.00000
51I	08/27/85	6	76	200	0	0.00000
51J	10/03/85	6	79	200	0	0.00000
61A	10/30/85	6	75	200	2	0.33333
61B	11/26/85	6	76	200	0	0.00000
61C	01/12/86	6	58	200	1	0.16667
	.	.	31	.	.	.
	.	.	30	.	.	.
	.	.	35	.	.	.
	.	.	40	.	.	.
	.	.	45	.	.	.
	.	.	50	.	.	.
	.	.	55	.	.	.
	.	.	60	.	.	.
	.	.	65	.	.	.
	.	.	70	.	.	.
	.	.	75	.	.	.
	.	.	80	.	.	.

We can plot the predicted and observed values as shown below. A vertical reference line at 31° is used to highlight the conditions at the Challenger launch.

The graph is shown in Figure below. There's not much data at low temperatures (the confidence band is quite wide), but the predicted probability of failure is uncomfortably high. Would you take a ride on Challenger when the weather is cold?

In [9]:
```
proc sort data=results;
  by predict;
run;

data results;
  set results;
  obs = Malfunction / Count;
run;

proc gplot data=results;
  plot (obs predict lower upper) * Temperature /
       href=31 lhref=33
       overlay frame vaxis=axis1 vminor=1;
  symbol1 v=dot i=none c=blue h=2;
  symbol2 v=none i=spline c=black w=5;
  symbol3 v=none i=spline c=red l=33 r=2 w=3;
  axis1 label=(a=90 "Estimated Failure probability") offset=(3);
run;
```

Out[9]: