STATS 604
Lecture 9

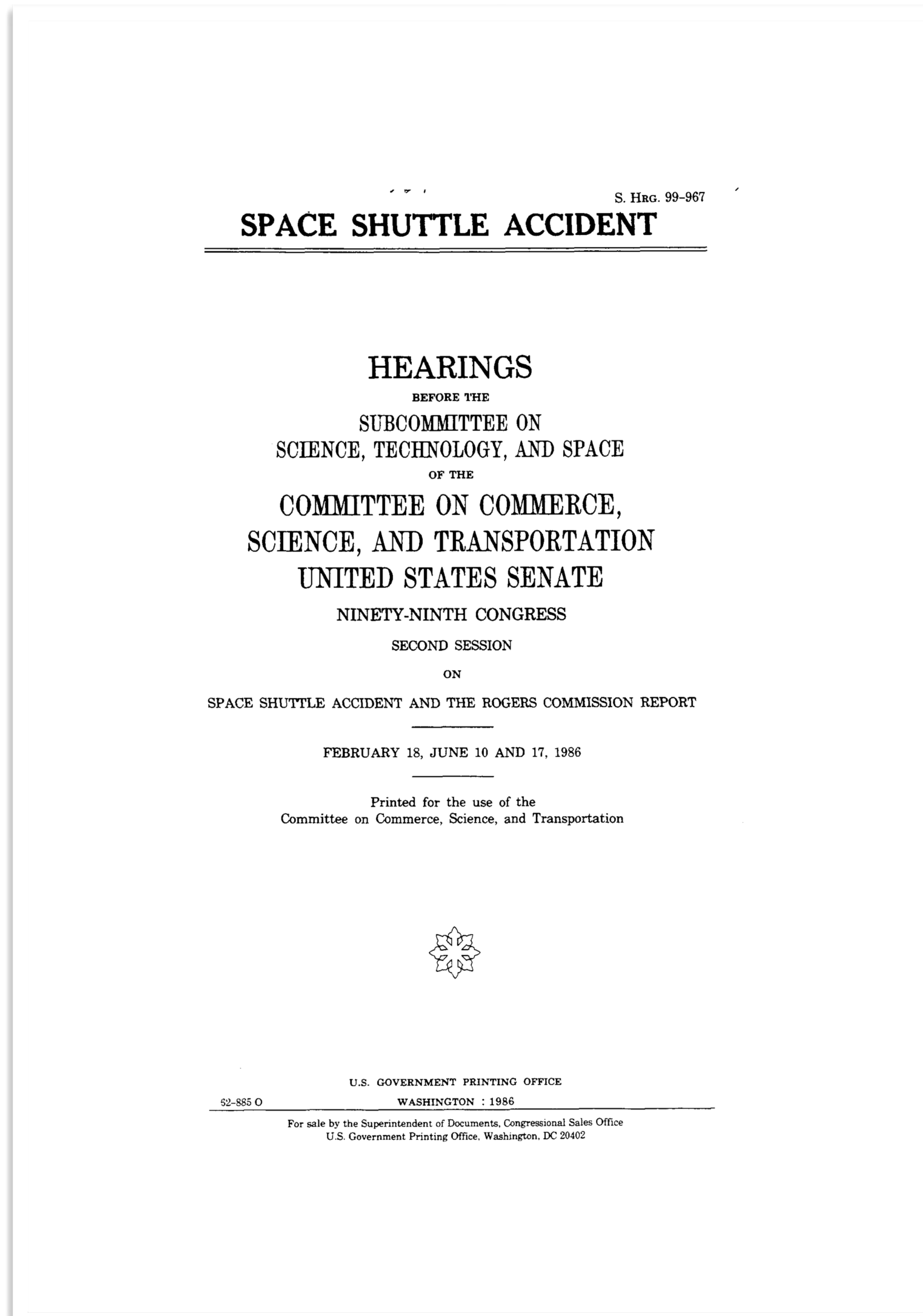
Quantifying uncertainty

Jonathan Terhorst

The Challenger Disaster

The Rogers Commission

- After the disaster, President Reagan formed a commission to investigate the causes.
- Notable members: Neil Armstrong, Sally Ride, Chuck Yeager, Richard Feynman.



S. HRG. 99-967

SPACE SHUTTLE ACCIDENT

HEARINGS

BEFORE THE

SUBCOMMITTEE ON

SCIENCE, TECHNOLOGY, AND SPACE

OF THE

COMMITTEE ON COMMERCE,

SCIENCE, AND TRANSPORTATION

UNITED STATES SENATE

NINETY-NINTH CONGRESS

SECOND SESSION

ON

SPACE SHUTTLE ACCIDENT AND THE ROGERS COMMISSION REPORT

FEBRUARY 18, JUNE 10 AND 17, 1986

Printed for the use of the

Committee on Commerce, Science, and Transportation

U.S. GOVERNMENT PRINTING OFFICE

52-885 O WASHINGTON : 1986

For sale by the Superintendent of Documents, Congressional Sales Office

U.S. Government Printing Office, Washington, DC 20402

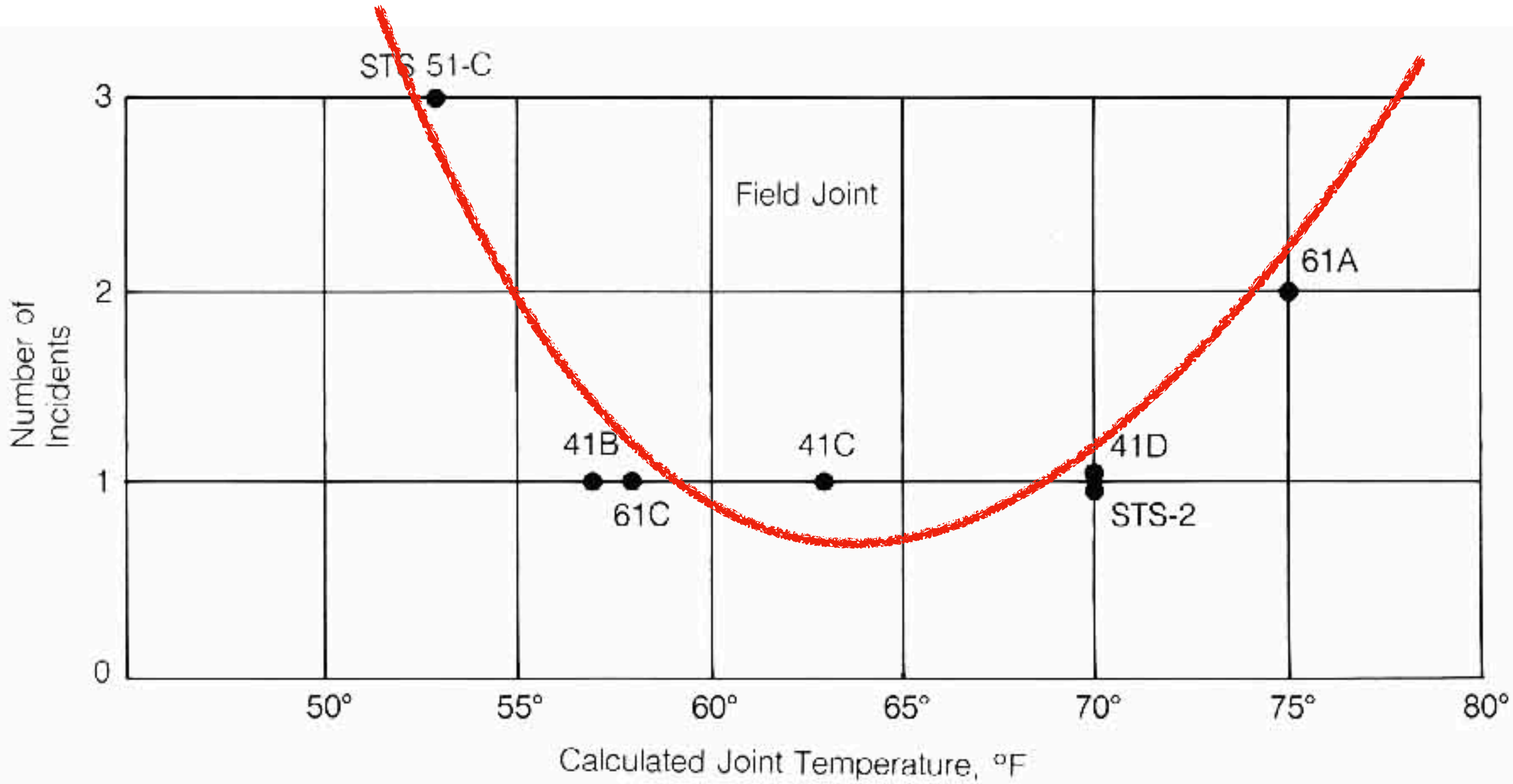


Figure 6
Plot of flights with incidents of O-ring thermal distress as function of temperature

"Based on the U configuration of points (identified by the flight number), it was concluded that there was no evidence from the historical data about a temperature effect."

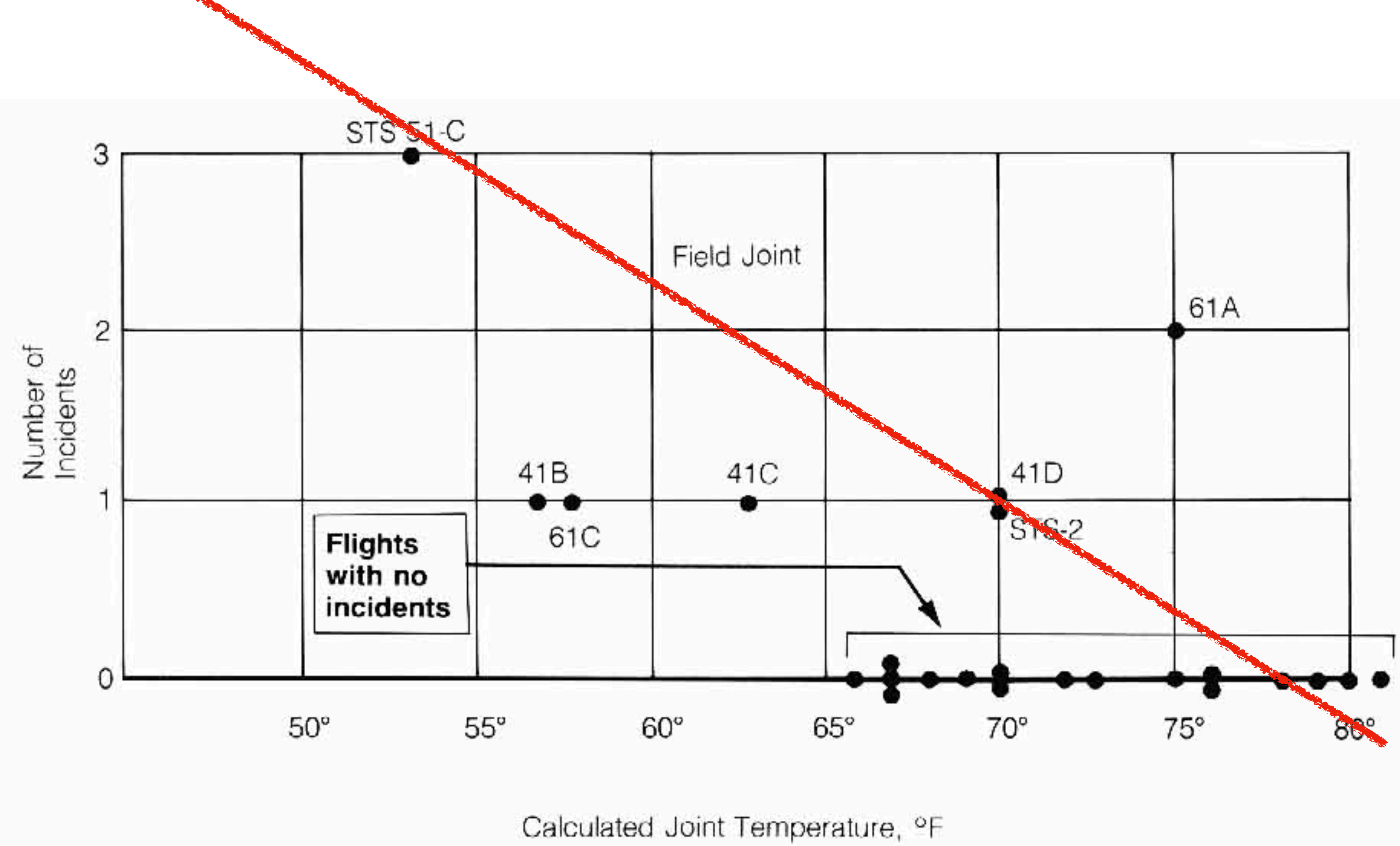


Figure 7
Plot of flights with and without incidents of O-ring thermal distress

"The Rogers Commission noted a mistake … the flights with zero incidents were left off the plot because it was felt that these flights did not contribute any information about the temperature effect."

The role of statistics

- Could statistical modeling have saved the Challenger?
- What are we trying to model?
- What are the data?

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

SIDDHARTHA R. DALAL, EDWARD B. FOWLKES, and BRUCE HOADLEY*

The Rogers Commission report on the space shuttle *Challenger* accident concluded that the accident was caused by a combustion gas leak through a joint in one of the booster rockets, which was sealed by a device called an O-ring. The commission further concluded that O-rings do not seal properly at low temperatures. In this article, data from the 23 preaccident launches of the space shuttle is used to predict O-ring performance under the *Challenger* launch conditions and relate it to the catastrophic failure of the shuttle. Analyses via binomial and binary logistic regression show that there is strong statistical evidence of a temperature effect on incidents of O-ring thermal distress. In addition, a probabilistic risk assessment at 31°F, the temperature at which *Challenger* was launched, yields at least a 13% probability of catastrophic field-joint O-ring failure. Postponement to 60°F would have reduced the probability to at least 2%. To assess uncertainty in estimates and for any future prediction under the *Challenger* scenario, a postanalysis prior distribution of the probability of a catastrophic failure is derived. The mean and median for this distribution for 31°F are at least .16 and .13, and for 60°F they are at least .004 and .02, respectively. The analysis of this article demonstrates that statistical science can play an important role in the space-shuttle risk-management process.

KEY WORDS: Catastrophic failure; Data analysis; O-rings; Probability risk assessment; Statistical science.

1. INTRODUCTION AND SUMMARY

On the night of January 27, 1986, the night before the space shuttle *Challenger* accident, there was a three-hour teleconference among people at Morton Thiokol (manufacturer of the solid rocket motor), Marshall Space Flight Center [NASA (National Aeronautics and Space Administration) center for motor design control], and Kennedy Space Center. The discussion focused on the forecast of a 31°F temperature for launch time the next morning, and the effect of low temperature on O-ring performance. A data set, Figure 1a, played an important role in the discussion. Each plotted point represents a shuttle flight that experienced thermal distress on the field-joint O-rings; the X axis shows the joint temperature at launch and the Y axis shows the number of O-rings that experienced some thermal distress. The O-rings seal the field joints of the solid rocket motors, which boost the shuttle into orbit. Based on the U configuration of points (identified by the flight number), it was concluded that there was no evidence from the historical data about a temperature effect.

Nevertheless, there was a debate on this issue, and some participants recommended that the launch be postponed until the temperature rose above 53°F—the lowest temperature experienced in previous launches—because the corresponding flight had the highest number of distressed O-rings. Some participants believed, based on the physical evidence, that there was a temperature effect on O-ring performance; for example, one of the participants, Roger Boisjoly, stated: "temperature was indeed a discriminator." In spite of this, the final recommendation of Morton Thiokol was to launch the *Challenger* on schedule. The recommendation transmitted to NASA stated that "Temperature data [are] not conclusive on predicting primary O-ring blowby." The same telefax stated that "Colder O-rings will have increased effective durometer ('harder'), and 'Harder' O-rings will take longer to 'seat'" [Presidential Commission Report, Vol. 1 (PC1), p. 97 (Presidential Commission on the Space Shuttle *Challenger* Accident 1986)].

After the accident a commission was appointed by President R. Reagan to find the cause. The commission was headed by former Secretary of State William Rogers and included some of the most respected names in the scientific and space communities. The commission determined the cause of the accident to be the following: "A combustion gas leak through the right Solid Rocket Motor aft field joint initiated at or shortly after ignition eventually weakened and/or penetrated the External Tank initiating vehicle structural breakup and loss of the Space Shuttle *Challenger* during mission 51-L" (PC1, p. 70). This is the type of failure that was debated the night before the *Challenger* accident.

The Rogers Commission (PC1, p. 145) noted that a mistake in the analysis of the thermal-distress data (Fig. 1a) was that the flights with zero incidents were left off the plot because it was felt that these flights did not contribute any information about the temperature effect (see Fig. 1b). The Rogers Commission concluded that "A careful analysis of the flight history of O-ring performance would have revealed the correlation of O-ring damage in low temperature" (PC1, p. 148).

This article aims to give more substance to this quote and show how statistical science could have provided valuable input to the launch decision process. Clearly, the key question was What would have constituted proof that it was unsafe to launch? Since the phenomenon is sto-

* Siddhartha R. Dalal is District Manager, Statistics and Econometrics Research Group, and the late Edward B. Fowlkes was Distinguished Member of Professional Staff, Statistics Research Group, Bellcore, Morristown, NJ 07960. Bruce Hoadley is Division Manager, Analytical Methods and Software Division, Navesink Research and Engineering Center, Red Bank, NJ 07701. The authors thank R. Gnanadesikan and J. R. Kettenring for encouraging them to work on this problem, and D. Duffy, J. Landwehr, C. Mallows, the editor, an associate editor, and the referees for helpful comments. This article is dedicated to the memory of Edward B. Fowlkes.

Journal of the American Statistical Association
December 1989, Vol. 84, No. 408, Applications & Case Studies

Modeling O-ring failure

- There are six O-rings per booster.
- Let t denote temperature and s denote pressure.
  - Let $p(s, t)$ denote the probability of an O-ring experiencing thermal distress.
- The number of failures per launch $X \mid s, t \sim \text{Binomial}(6, p(s, t))$

Logistic regression model

$$\log\left(\frac{p(s,t)}{1-p(s,t)}\right) = \alpha + \beta t + \gamma s.$$

```
Call:
glm(formula = damaged/6 ~ temp + pressure, family = "binomial",
    data = orings, weights = rep(6, nrow(orings)))

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.0361  -0.6434  -0.5308  -0.1625   2.3418

Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept)  2.520195   3.486784   0.723   0.4698
temp        -0.098297   0.044890  -2.190   0.0285 *
pressure     0.008484   0.007677   1.105   0.2691
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 24.230  on 22  degrees of freedom
Residual deviance: 16.546  on 20  degrees of freedom
AIC: 36.106
```

Residual deviance

- From the last slide:

    Null deviance: 24.230 on 22 degrees of freedom
Residual deviance: 16.546 on 20 degrees of freedom
AIC: 36.106

- Recall that
    - Null deviance: 2*[LL(Saturated model) - LL(null model)]
    - Residual deviance: 2*[LL(Saturated model) - LL(the model)]
- If the model is good, residual deviance is low.

Does pressure matter?

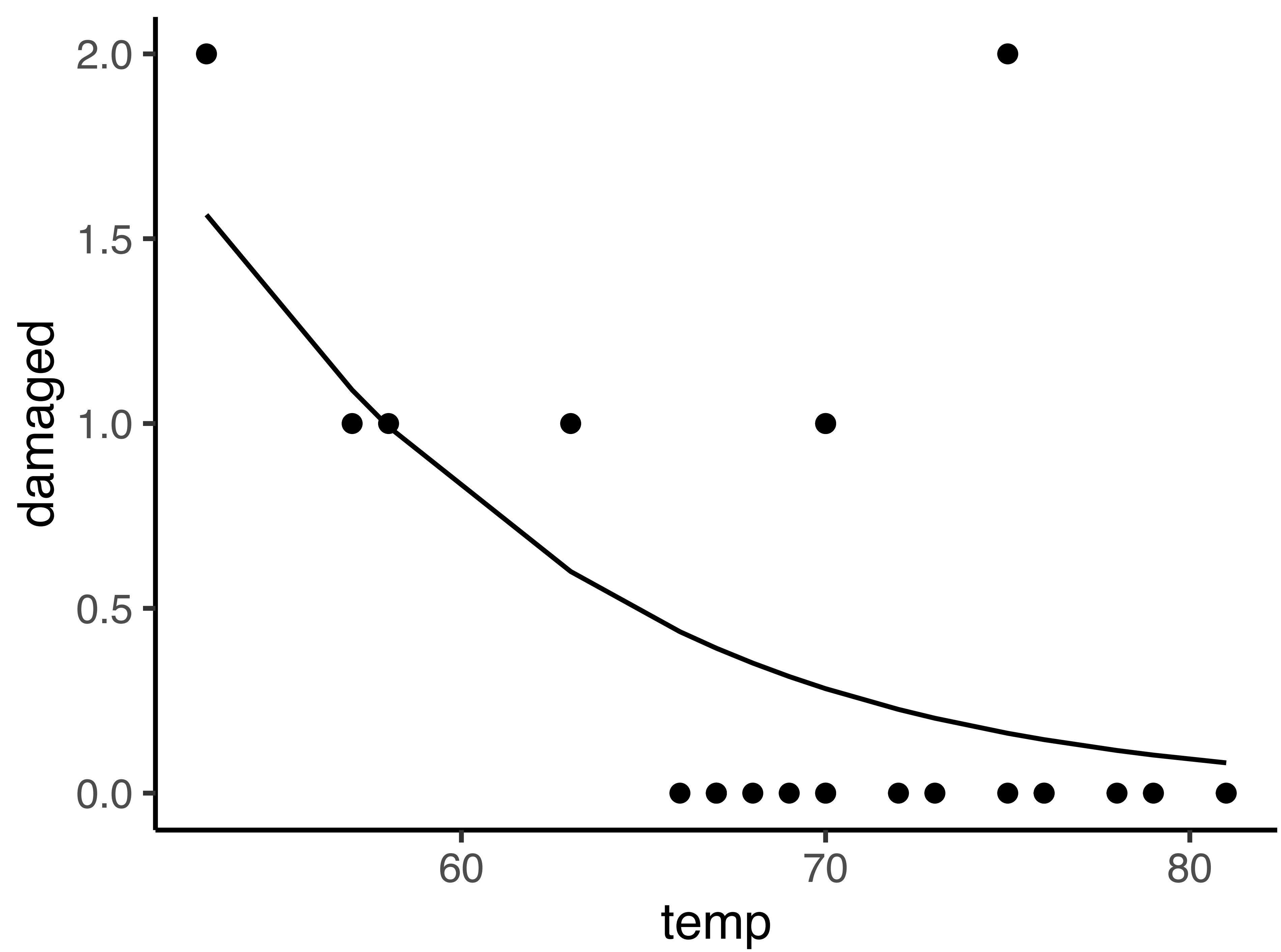
- “90% bootstrap confidence intervals for the expected numbers of incidents were calculated for each temperature, first holding pressure constant at 50 psi and next setting the pressure at 200 psi.”

- **Exercise**: what are the 90% bootstrap CI’s for the expected number of incidents at $s = 50$ and $s = 200$?

Pressure-only model

```
Call:
glm(formula = damaged/6 ~ temp, family = "binomial", data = orings,
    weights = rep(6, nrow(orings)))

Deviance Residuals:
     Min        1Q    Median        3Q       Max
-0.95227  -0.78299  -0.54117  -0.04379   2.65152

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  5.08498    3.05247   1.666   0.0957 .
temp        -0.11560    0.04702  -2.458   0.0140 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 24.230  on 22  degrees of freedom
Residual deviance: 18.086  on 21  degrees of freedom
AIC: 35.647

Number of Fisher Scoring iterations: 5
```

Predicted vs. observed

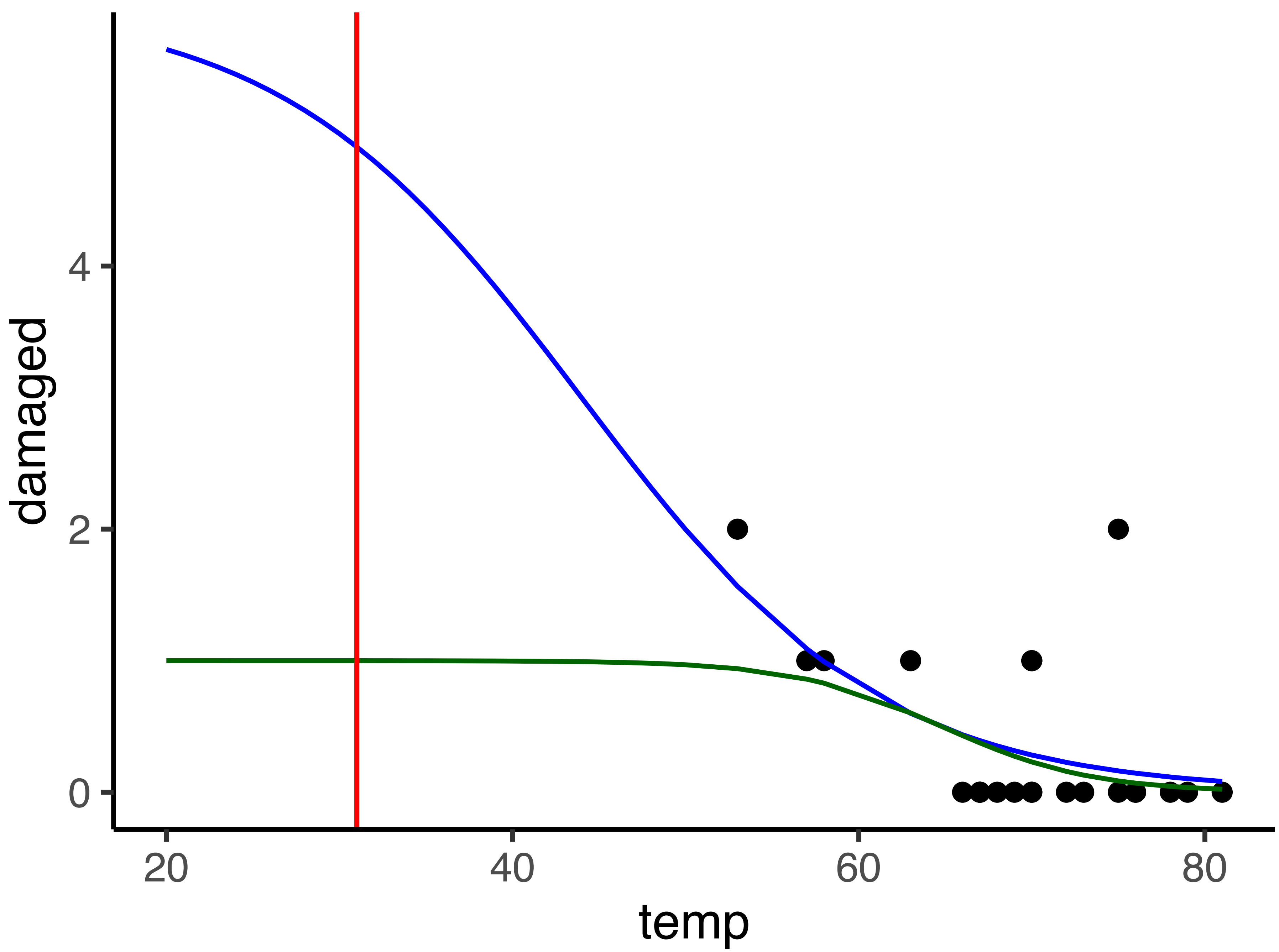


Alternative model

- The prediction of interest is really whether *any* O-rings will fail.

```
Call:
glm(formula = pmin(damaged, 1) ~ temp, family = "binomial", data = orings)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.0611  -0.7613  -0.3783   0.4524   2.2175

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  15.0429     7.3786   2.039   0.0415 *
temp        -0.2322     0.1082  -2.145   0.0320 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 28.267  on 22  degrees of freedom
Residual deviance: 20.315  on 21  degrees of freedom
AIC: 24.315

Number of Fisher Scoring iterations: 5
```

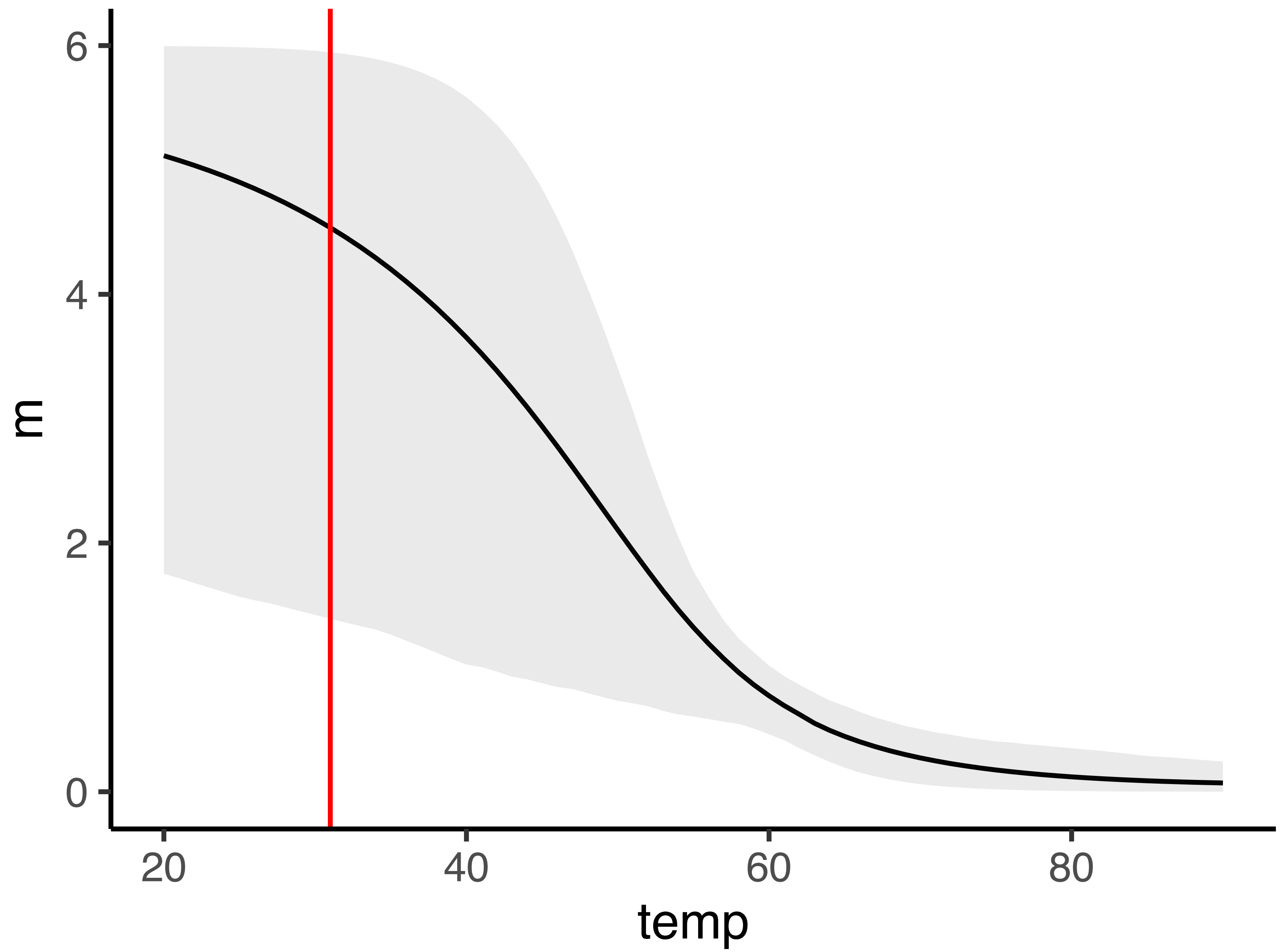
Binomial vs. logit model



Confidence bands

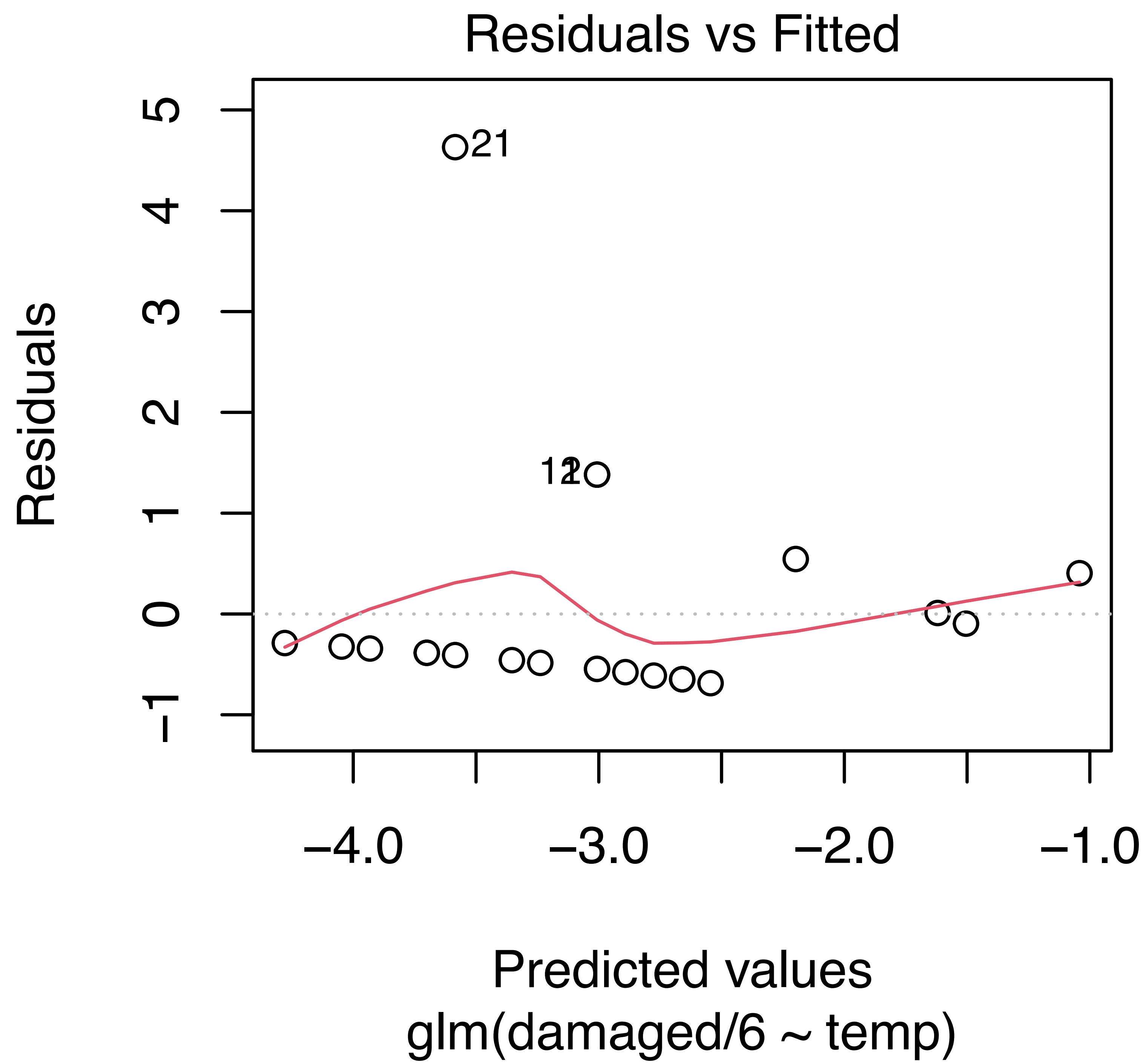
- Both models predict that at 31F, about 4-5 O-rings would fail.
- How confident can we be in this prediction?
- We will generate (pointwise) confidence bands using the parametric bootstrap.

Bootstrap confidence bands



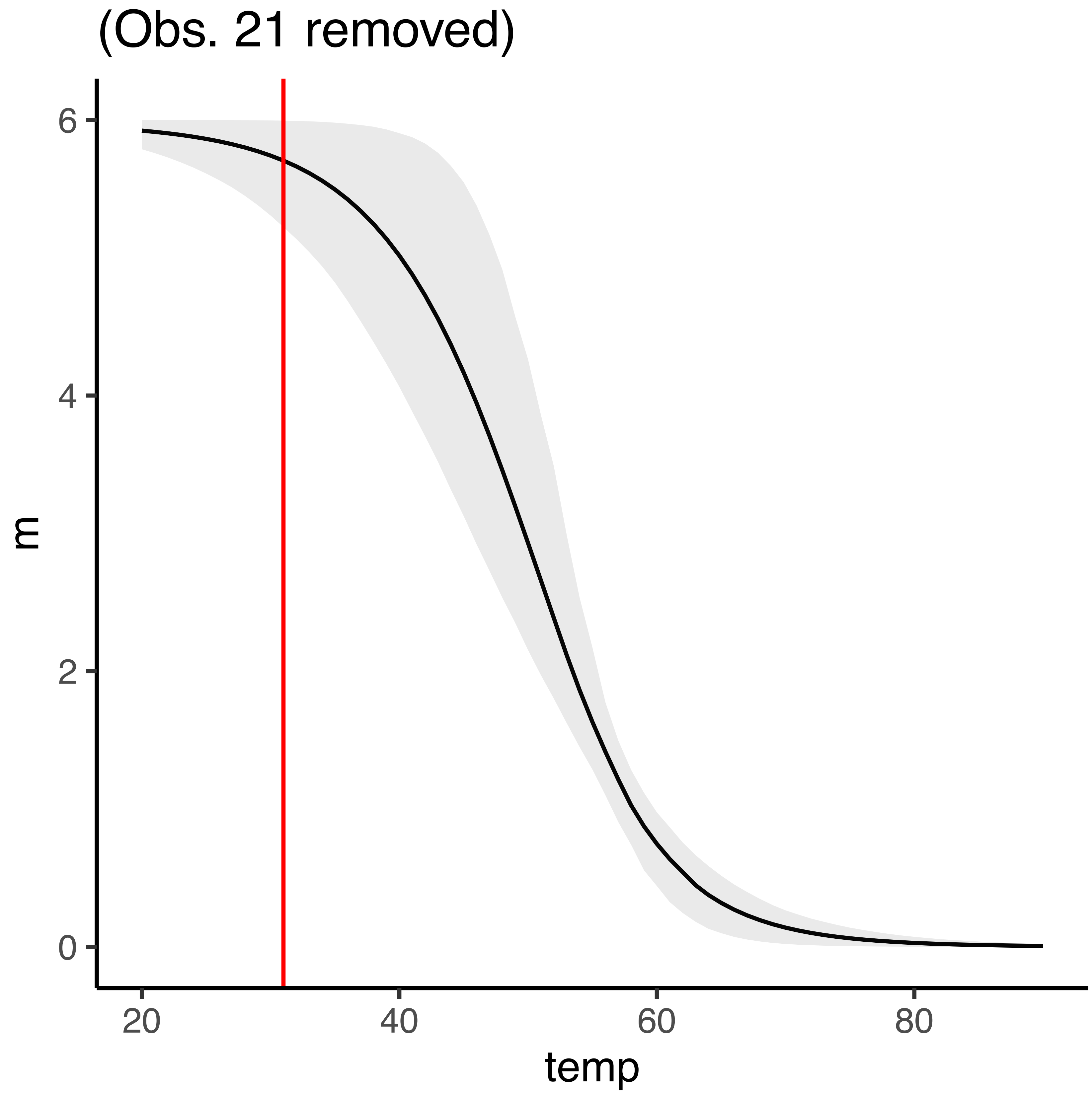
Stability checking

- Are there any outliers? Are our findings robust to them?



Removing outlying observation 21

```
glm(formula = damaged/6 ~ temp, family = "binomial", data = orings,
    weights = rep(6, nrow(orings)), subset = order != 21)

Deviance Residuals:
    Min        1Q    Median        3Q       Max
-0.7608   -0.5742   -0.3321   -0.1861    1.5204

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  8.66157    3.63441   2.383  0.01716 *
temp        -0.17680    0.05869  -3.013  0.00259 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 20.0667  on 21  degrees of freedom
Residual deviance:  9.4096  on 20  degrees of freedom
AIC: 24.748

Number of Fisher Scoring iterations: 6
```


(Obs. 21 removed)
m
temp
0
2
4
6
20
40
60
80