Center for Machine Learning and Intelligent Systems

Challenger USA Space Shuttle O-Ring Data Set

Download: Data Folder, Data Set Description

Abstract: Task: predict the number of O-rings that experience thermal distress on a flight at 31 degrees F given data on the previous 23 shuttle flights

Data Set Characteristics:	Multivariate	Number of Instances:	23	Area:	Physical
Attribute Characteristics:	Integer	Number of Attributes:	4	Date Donated	1993-08-05
Associated Tasks:	Regression	Missing Values?	No	Number of Web Hits:	127458

Source:

Original Owner:

David Draper (draper '@' math.ucla.edu)
University of California, Los Angeles

Donor:

David Draper (draper '@' math.ucla.edu)

Data Set Information:

There are two databases: (both use the same set of 5 attributes):
1. Primary o-ring erosion and/or blowby
2. Primary o-ring erosion only

The two databases are identical except for the 2nd attribute of the 21st instance (confirmed by David Draper on 8/5/93).

Edited from (Draper, 1993):

The motivation for collecting this database was the explosion of the USA Space Shuttle Challenger on 28 January, 1986. An investigation ensued into the reliability of the shuttle's propulsion system. The explosion was eventually traced to the failure of one of the three field joints on one of the two solid booster rockets. Each of these six field joints includes two O-rings, designated as primary and secondary, which fail when phenomena called erosion and blowby both occur.

The night before the launch a decision had to be made regarding launch safety. The discussion among engineers and managers leading to this decision included concern that the probability of failure of the O-rings depended on the temperature t at launch, which was forecase to be 31 degrees F. There are strong engineering reasons based on the composition of O-rings to support the judgment that failure probability may rise monotonically as temperature drops. One other variable, the pressure s at which safety testing for field join leaks was performed, was available, but its relevance to

the failure process was unclear.

Draper's paper includes a menacing figure graphing the number of field joints experiencing stress vs. liftoff temperature for the 23 shuttle flights previous to the Challenger disaster. No previous liftoff temperature was under 53 degrees F. Although tremendous extrapolation must be done from the given data to assess risk at 31 degrees F, it is obvious even to the layman "to foresee the unacceptably high risk created by launching at 31 degrees F." For more information, see Draper (1993) or the other previous analyses.

The task is to predict the number of O-rings that will experience thermal distress for a given flight when the launch temperature is below freezing.

Attribute Information:

1. Number of O-rings at risk on a given flight
2. Number experiencing thermal distress
3. Launch temperature (degrees F)
4. Leak-check pressure (psi)
5. Temporal order of flight

Relevant Papers:

Draper,D. (1993). Assessment and propagation of model uncertainty. In Proceedings of the Fourth International Workshop on Artificial Intelligence and Statistics (pp. 497--509). Ft. Lauderdale, FL: Unpublished.
[Web Link]

Dalal,S.R., Fowlkes,E.B., & Hoadley,B. (1989). Risk analysis of the space shuttle: pre-Challenger prediction of failure. Journal of the American Statisticians Association, 84, 945--957.
[Web Link]

Lavine,M. (1991). Problems in extrapolation illustrated with space shuttle O-ring data. Journal of the American Statisticians Association, 86, 919--922.

Martz H.F., & Zimmer, W.J. (1992). The risk of catastrophic failure of the solid rocket boosters on the space shuttle. American Statistics, 46, 42--47.
[Web Link]

Papers That Cite This Data Set[1]:

Stephen D. Bay. Multivariate Discretization for Set Mining. Knowl. Inf. Syst, 3. 2001. [View Context].

Pedro Domingos. Linear-Time Rule Induction. KDD. 1996. [View Context].

Mohammed Waleed Kadous and Claude Sammut. The University of New South Wales School of Computer Science and Engineering Temporal Classification: Extending the Classification Paradigm to Multivariate Time Series. [View Context].

Citation Request:

Please refer to the Machine Learning Repository's citation policy

---