Deloitte.

Testing Benford's Law with Software Code Counts

Chris Kaldes
Chuck Knight

ICEAA Conference
June 2014

Agenda

- Introduction
- What is Benford's Law?
- Situation and Problem
- Testing the law with code counts
- Practical applications for cost estimating
- Questions

Deloitte.

Introduction

Chris Kaldes

CCE/A, PMP

ckaldes@deloitte.com

20 years of experience in cost estimating and decision analysis including many cost studies and investment determinations for government and private sector decision-makers.

Currently supporting USPS to develop should-costs for software development projects

Chuck Knight

CCE/A

chuknight@deloitte.com

5 years with Deloitte and Dept. of the Navy in cost estimating, earned value management, and general acquisition management

Project concentration within Intel Community and software development cost estimating

Deloitte.

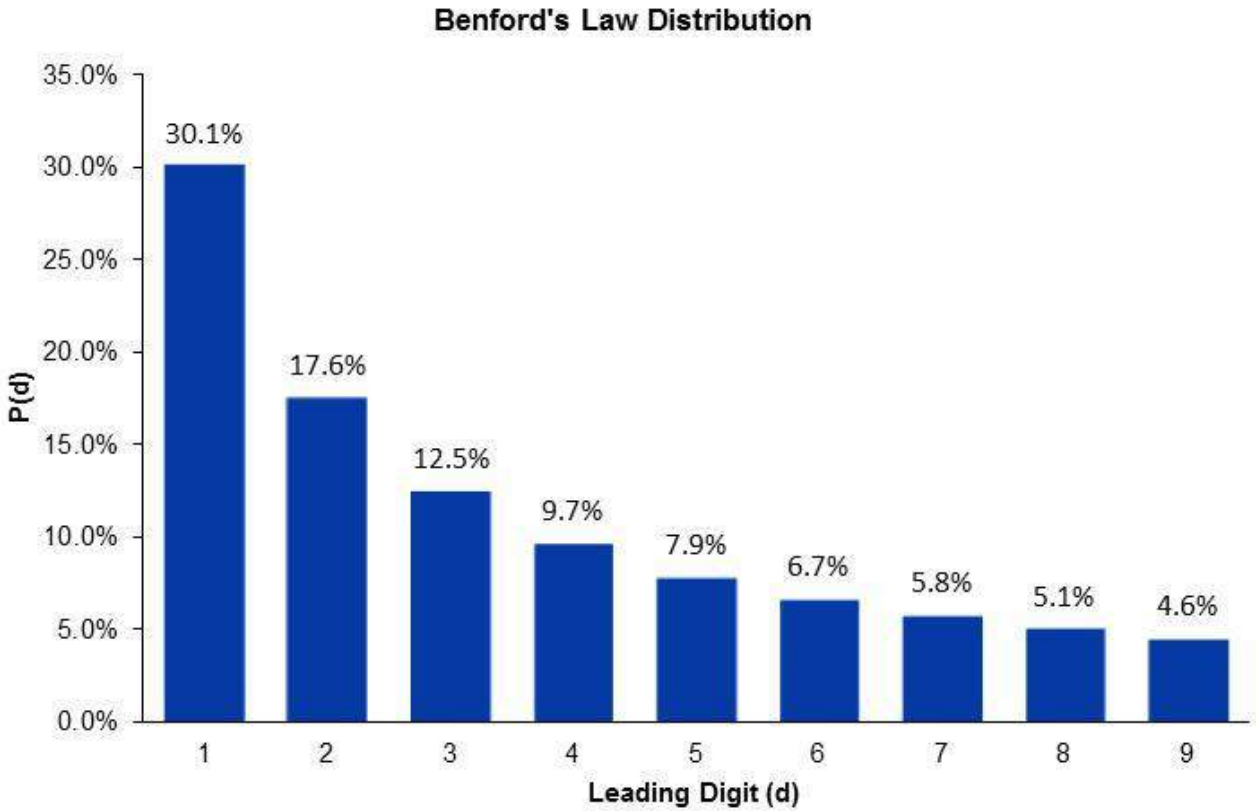
Basics of Benford's Law

Benford's Law looks at the leading digit of each data point in any real data set. The law simply states to then look at the frequency of each leading digit. In other words, how many data points start with one? How many start with two? And so on.

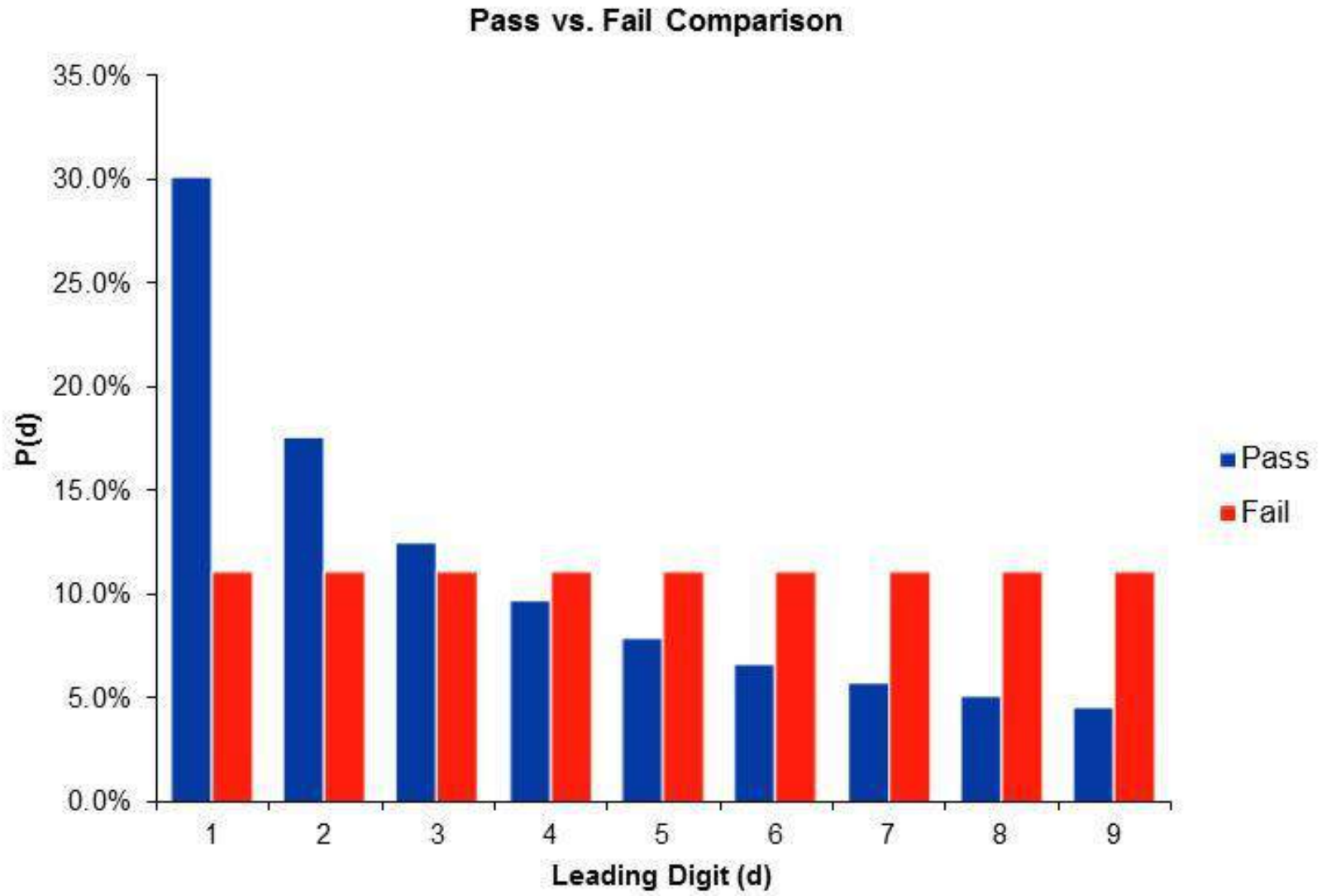
If you were to create a histogram of the frequency of each leading digit (one through nine), what pattern, or distribution, do you think you would see? (e.g. uniform, normal)

Basics of Benford's Law

Given any real data set, Benford's Law states that the leading digit of each data point should exhibit a distinct frequency pattern where the number one occurs more often than the number two which occurs more often than the number three and so on.



Benford's Law: Pass vs. Fail Comparison



How can we use Benford's Law as cost estimators?

Situation: Tasked to create a life cycle cost estimate (LCCE) for a ground system which will primarily consist of software development. The primary basis of estimate are source lines of code (SLOC) count estimates provided in the Cost Analysis Requirements Description (CARD). Assume the CARD contains best available data.

Problem: How confident are you in the SLOC estimates provided in the CARD? If confidence is low, what basis do we have to challenge the estimates?

Hypothesis: Actual SLOC counts (real data set based on software that has already been developed) will follow Benford's Law.

Question: If hypothesis is true, should the same logic be applied to the SLOC estimates from the CARD that are used as primary basis of the LCCE?

Deloitte.

Ground Rules and Assumptions

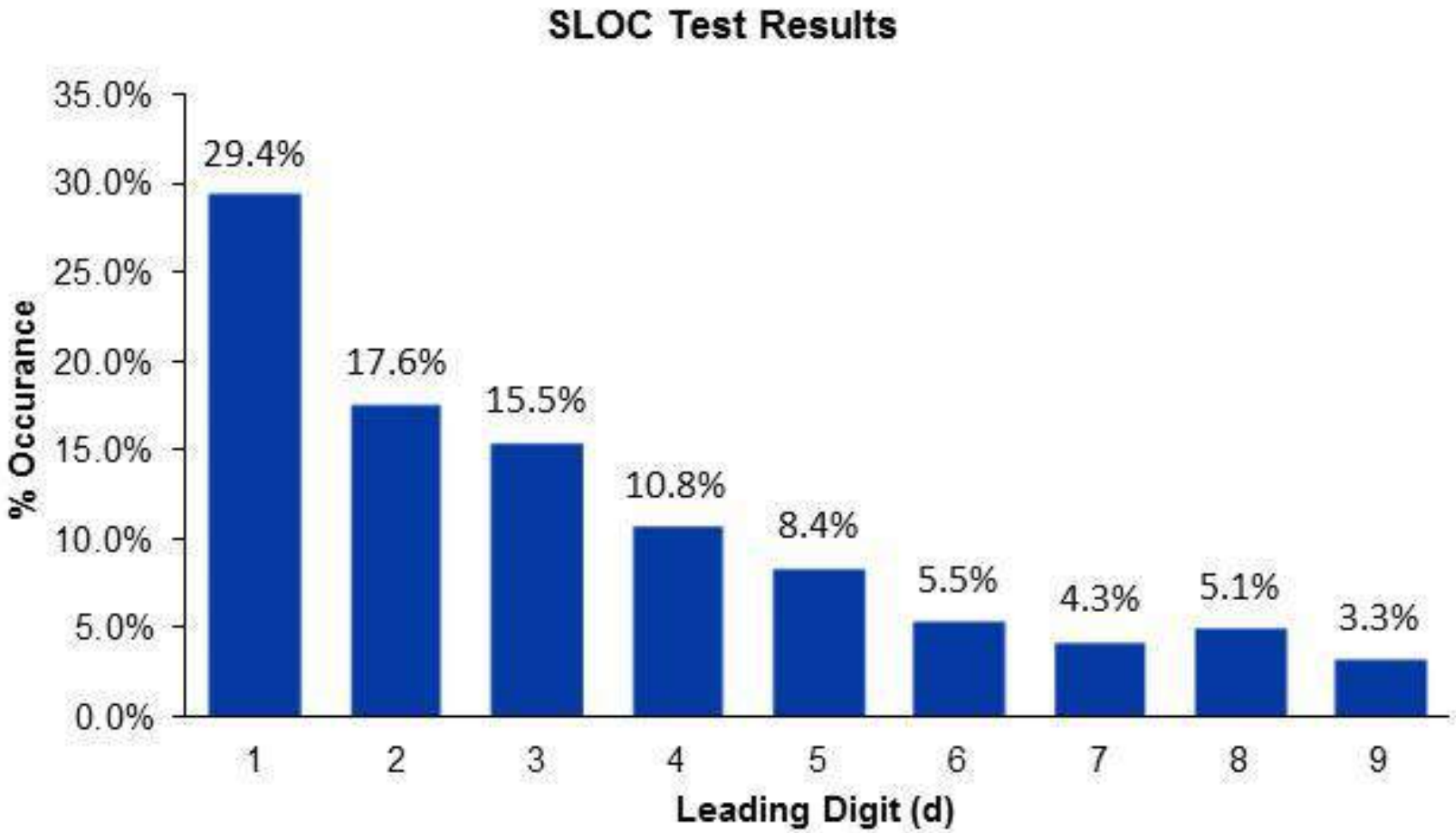
General:

- Data set must consist of real data
- Each data point is independent of each other
- Zero or non-value data points are excluded

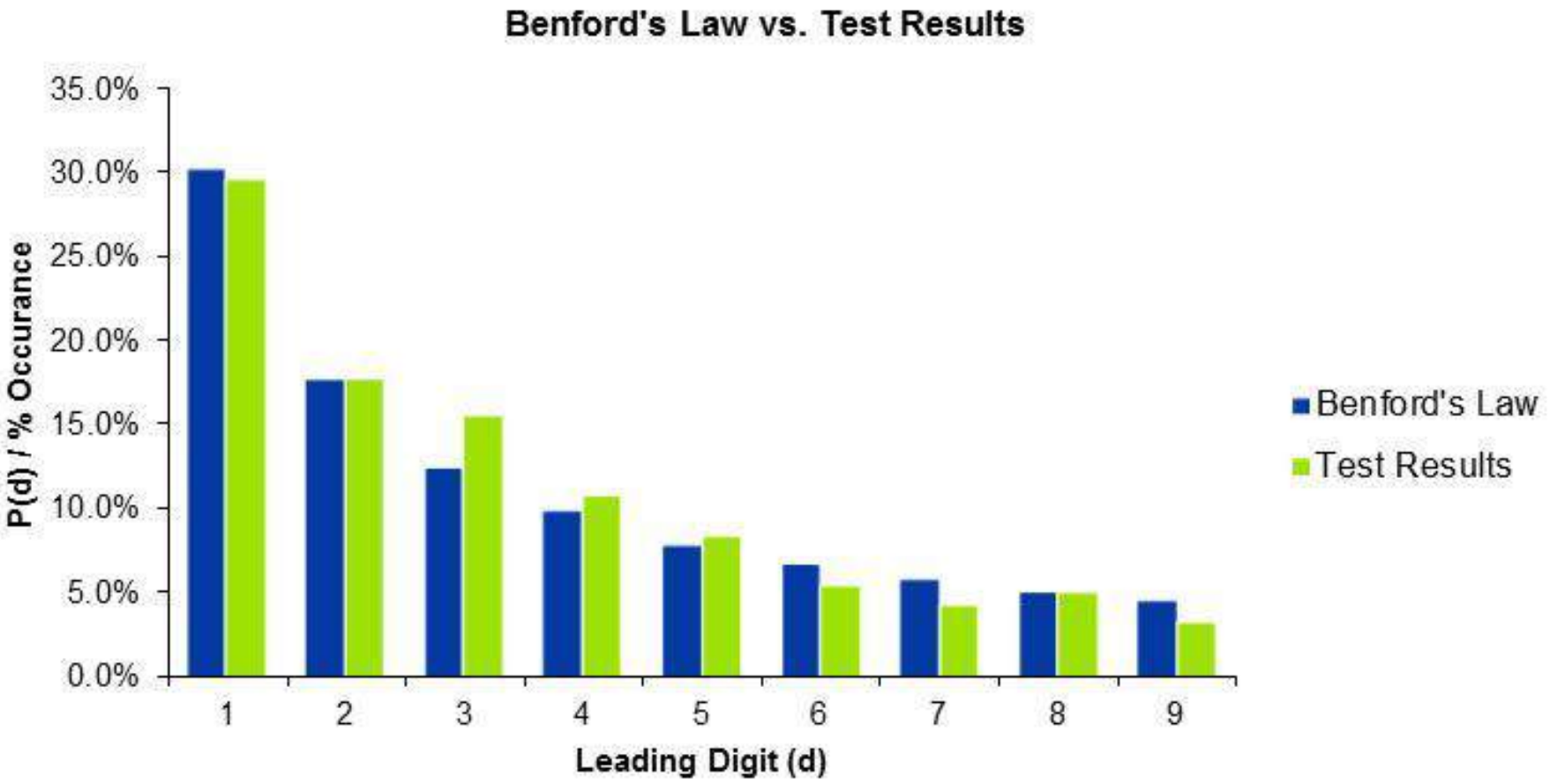
SLOC Test Specific:

- 510 data points (Actual SLOC counts from final delivered software)
- Counts are at lower levels (CSCI or CSC); not at a system or program level

Benford's Law Test Results



Test Results Comparison



Deloitte.

Conclusion

If we believe that the actual SLOC counts will follow the behavior consistent with Benford's Law, then we should stress that the SLOC estimates at the beginning of the LCCE development should also follow that behavior.

This test is meant to be a quick and easy cross check for cost estimators who may lack subject matter expertise in technical areas that they have been tasked to estimate.

If input data does not pass the Benford's Law test, it can provide the cost estimator a starting point to go back to the engineers or SME's and explore the basis of estimate for those inputs.

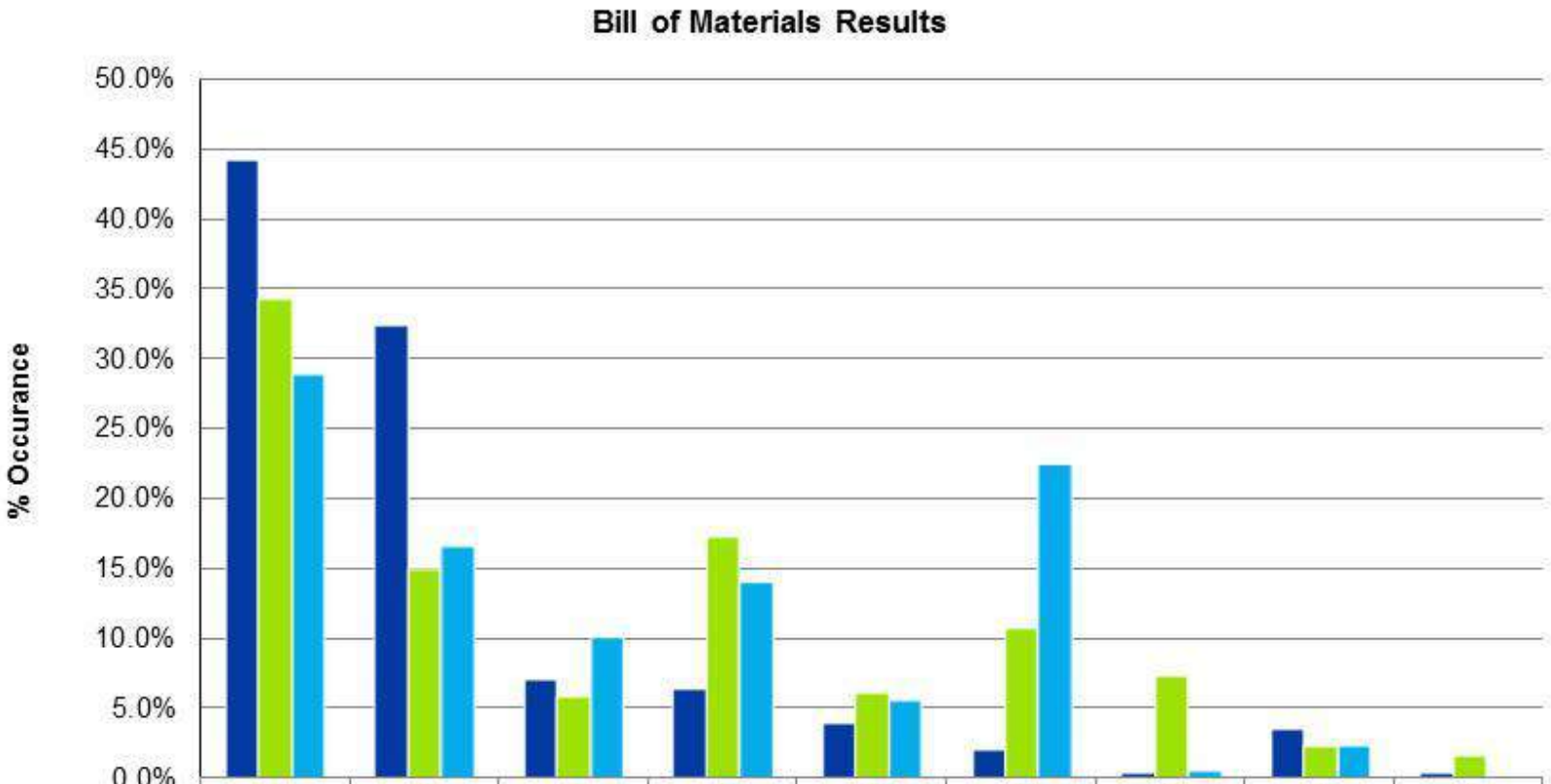
At the very least, this test can help provoke additional thought around inputs which will help make an estimate more defendable.

This test can help increase stakeholders' confidence in technical inputs

Deloitte.

Discussion: Practical Applications for Cost Estimating

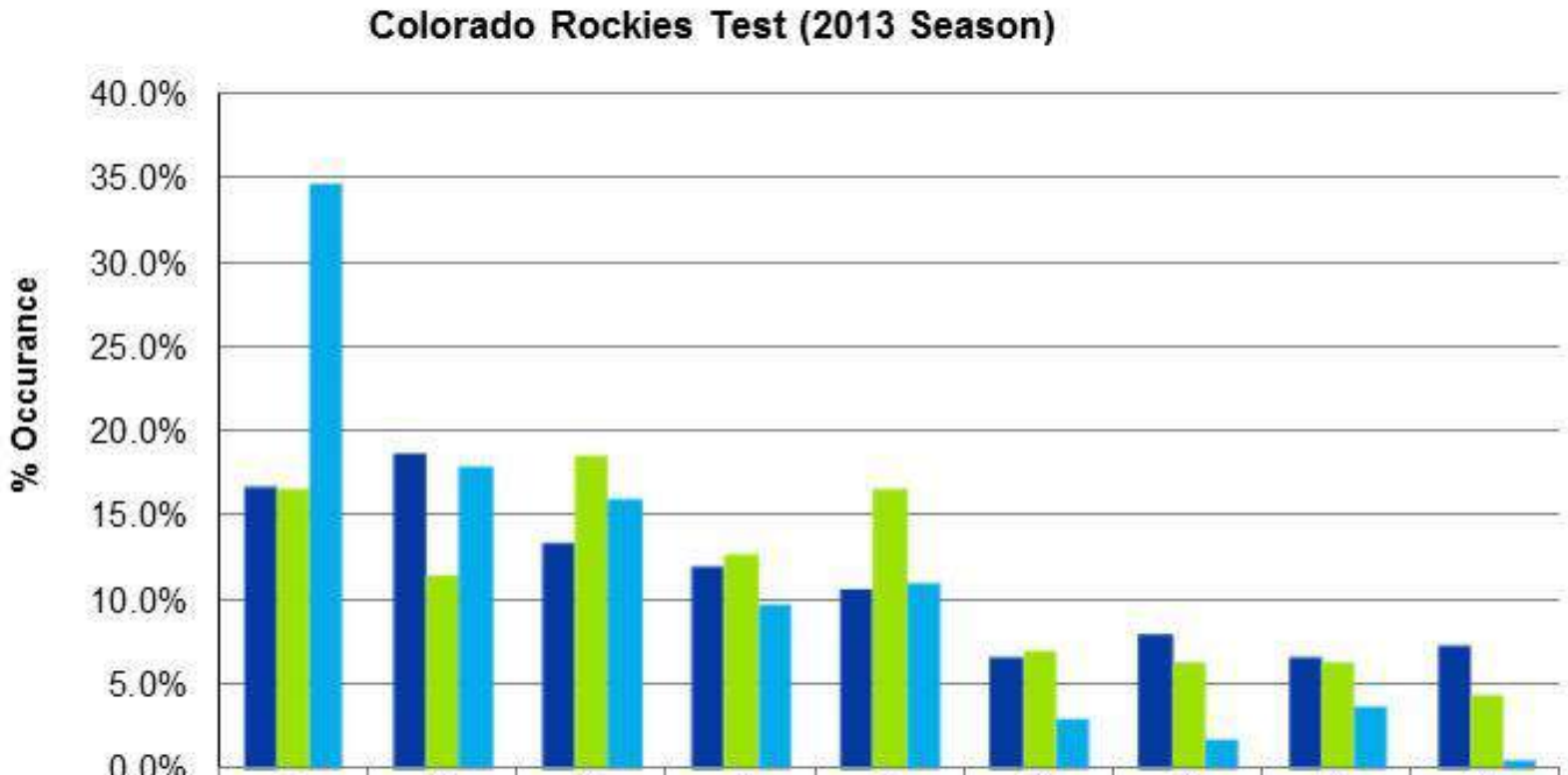
What other areas might we be able to use this test?



	1	2	3	4	5	6	7	8	9
Qty	44.2%	32.3%	7.0%	6.4%	3.9%	2.0%	0.4%	3.5%	0.4%
Unit Cost	34.2%	15.0%	5.8%	17.2%	6.0%	10.7%	7.4%	2.2%	1.6%
Maint Cost	28.8%	16.5%	10.1%	14.0%	5.6%	22.3%	0.6%	2.2%	0.0%

Deloitte.

Other Benford's Law Examples



	1	2	3	4	5	6	7	8	9
Runs Scored	16.7%	18.7%	13.3%	12.0%	10.7%	6.7%	8.0%	6.7%	7.3%
Runs Against	16.6%	11.5%	18.5%	12.7%	16.6%	7.0%	6.4%	6.4%	4.5%
Margin of Victory	34.6%	17.9%	16.0%	9.9%	11.1%	3.1%	1.9%	3.7%	0.6%

Deloitte.

Wrap Up

Questions?

Deloitte.