



# GLM: The Predictive Modeling Context With Minimum Bias, GLMs and Credibility

Moderator: Curtis Gary Dean, FCAS, MAAA  
Lincoln Distinguished Professor of Actuarial Science

---

# My Interest in Predictive Modeling

---

- 1989 article in *Science*
- “Clinical Versus Actuarial Judgment”
- Summarized in 1990 in *Contingencies*

# Clinical Versus Actuarial Judgment

---

- *“In the clinical method the decision-maker combines or processes information in his or her head.”*
- *“In the actuarial or statistical method the human judge is eliminated and conclusions rest solely on **empirically established relations** between data and the condition or event of interest.”*

# Most Recent Predictive Modeling Experience

---

- 2017 NCAA Basketball Tournament Bracket
- Supervised two students, Tim Hoblin and Cody Kocher, for their honors' thesis
- Wealth of sports data. Used data from 2006 thru 2016 to build and test models (11 years)
- Submitted their 2017 predictions to [espn.com](http://espn.com) for scoring and ranking against 18 million+ entries

# How did they do?

---

- Submitted 24 completed brackets: 22 based on “*empirically established relations*” and 2 based on “*clinical judgments*” as devoted fans
- Based on ESPN’s scoring algorithm:
  - 18 of their 22 modeled brackets were above the 50<sup>th</sup> percentile in the 18 million+ entries
  - Their “clinical judgments” did poorly: 22.7<sup>th</sup> and 18.9<sup>th</sup> percentiles
  - Most important: Cody submitted his strongest model in his family challenge and beat his sister for the first time in years

## *“empirically established relations”*

---

1. Created models using 8 years of data. Used GLM (logistic regression) and random forests.
  2. Validated models on 3 holdout years
  3. Measured predictive power on the 3 holdout years.
- There are a variety of tools to accomplish 1.
  - 2 and 3 should be done on out-of-sample data
  - **Warning:** Do not blindly assume that your GLM has predictive power!

# Tools to Do Predictive Modeling

---

- The new tools: GLM, Random Forests, Neural Networks, CART, MARS, etc.
- The established actuarial tools: Bailey and Simon's minimum bias procedure, age-to-age factor methods, credibility theory, etc.
- What is the cornerstone of predictive modeling?
  - *Not the particular tool that was used.*
  - *It's the validation of the model using data.*