

STATtr@k

A Statistician in an Insurance Agency

Andy Pulkstenis

So you've made the decision to pursue a career in statistics, but what now? Generally speaking, there are four primary employment avenues:

- Academia
- Government
- Biostatistics
- Business

Multiple articles could be written about each of those areas, but I want to speak specifically to those who are contemplating an applied statistics career in a business setting.

Why Should I Consider Business?

Security: Whether you call it business analytics, data mining, predictive modeling, or applied statistics (at State Farm, we call it "advanced analytics"), growth in this area has been explosive

over the past 10 years, and the outlook appears even more outstanding.

Various companies (and even entire industries) are realizing the importance of proper analytics and the value tools like predictive modeling, experimental design, and regression-based business analysis add to a firm. When they make a successful movie starring Brad Pitt about baseball teams using statistical analysis to direct player acquisition strategies, it's safe to declare that analytics has gone mainstream. The emergence of Big Data concepts continues to force the issue, requiring old school, Excel-based MBA approaches to be replaced with more powerful statistically based methods.

As demonstrated by "Moneyball," better analysis leads to better strategic insight, which in turn leads to better decisions and a direct positive impact on the metrics that matter to an organization. The statistician's toolbox comes loaded with methodology and approaches well suited to the problems encountered in a business setting, and quite frankly the demand continues to dramatically outpace supply. It was an issue when I entered the work force nearly 20 years ago, and it's even worse today. If you are a well-trained statistician, that's what we call a "buyer's market." At State Farm, we're always on the lookout for quality graduate-level statisticians.

Variety: People often comment that you get to play in everybody else's sandbox as a statistician, and that is true in business settings as well. Throughout my career I've worked on problems involving credit cards, loans, marketing, the Internet, greeting cards, grocery stores, airlines, tire manufacturers, auto manufacturers, theme parks, chemical manufacturers, and pharmaceutical companies. In my current role at State Farm, my analytic project partners have included actuaries, our banking business, marketing, our Internet group, insurance agents, claims, IT....It's tough to get bored when you are always exposed to new parts of the business.

Variety (Part II): So the business areas are varied, but what about the methodology? Many people have the mistaken idea that applied statistics in a business setting is basically straightforward logistic regression, linear regression, confidence intervals, rinse and repeat, but that couldn't be further from the truth. Sure, those techniques are widely used and applicable to a broad number of business problems, but I've also been involved with projects using advanced time series methodology, cluster analysis, PCA, nonlinear regression, experimental design,

What gets you the job?

- Master's degree or above (statistics)
- Internships/applied experience
- Communication of technical concepts, especially the ability to map a business problem into the statistical world and map the stat solution back into business terminology
- Problemsolving
- Technical ability (know "why" vs. just knowing "how")
- Logistic regression with Big Data
- GLM/regression theory
- Breadth and depth beyond logistic and GLM
- Useful "niche" skills like DOE, time series, survival analysis, or multivariate methods

Deal breakers?

- Poor communication
- Technical *inability*
- Being wrong, but convinced you are right
- Lack of problemsolving ability
- Faking it (know what's on your résumé)

machine learning, multivariate outlier detection, decision trees, and survival analysis over the past year. One thing I love about business analytics is how often and deeply we reach into the corners of the stat toolkit.

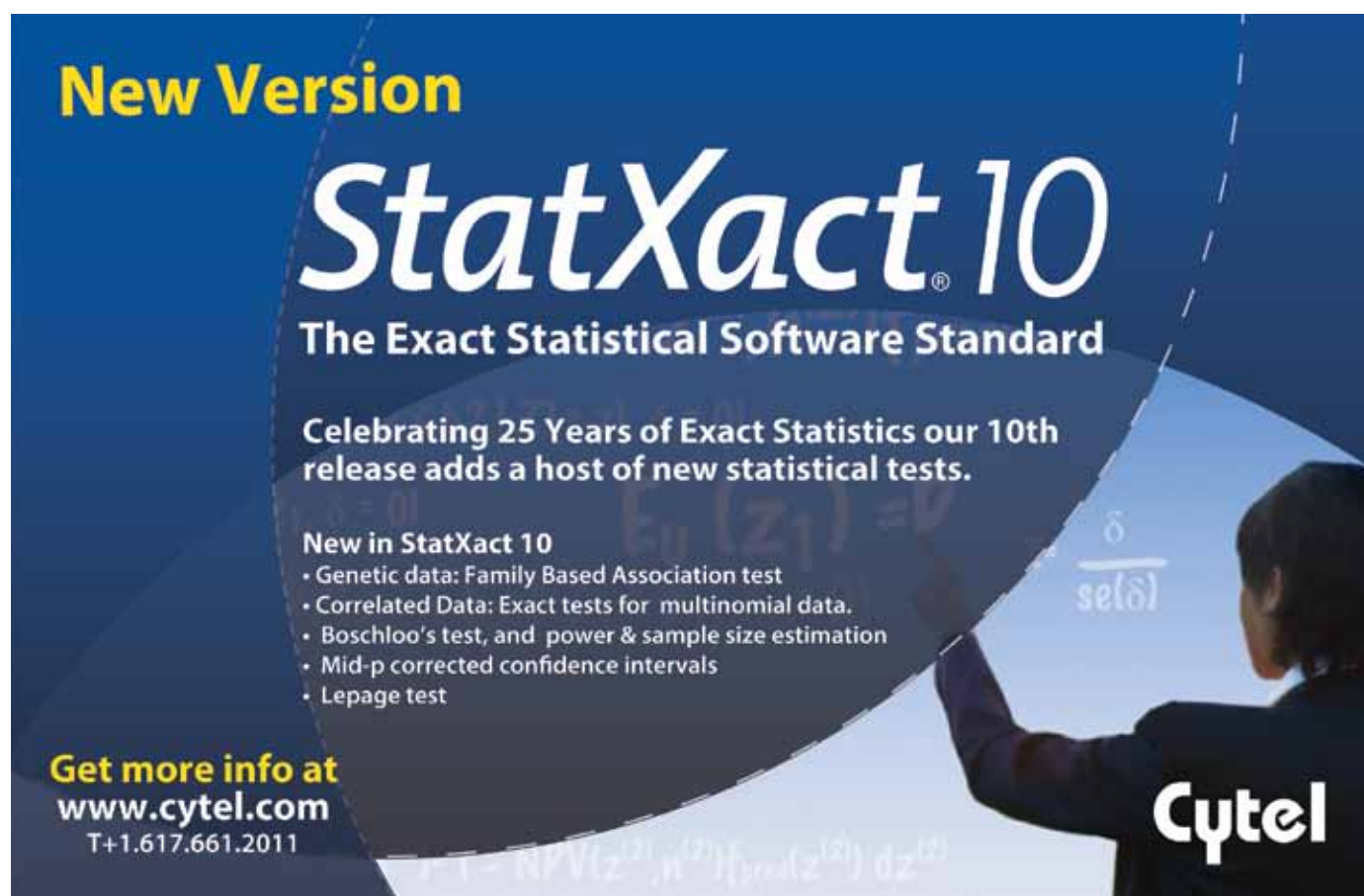
Challenge: The complexity of the problems we face in business is primarily driven by the following:

1. The size and inherent bias of internal observational data sources
2. Regulatory or organizational constraints around any solution we'd like to implement
3. An environment in which many conditions theoretically required for clean methodology are violated in one way or another

The combination of all three often leads to challenging situations in which we are trying to find the most helpful yet "least wrong" approach to the problem versus finding the theoretically correct answer. And because business is competitive, there

is not usually a ton of published research to draw from. It may require pulling analysis ideas from biostatistics to apply to a small sample problem in auto research, or applying manufacturing quality control ideas to some non-manufacturing business process. Recently, my team had to find a way to apply predictive modeling concepts in a useful way to a situation in which we didn't have a single instance of a target variable! It can be quite challenging, but incredibly rewarding, when successfully solving some previously "unsolvable" problem that doesn't quite fit into anything we learned in our books.

Academia, government, and biostatistics tend to get more promotion at the university level, but don't neglect this growing and rewarding field of "advanced analytics" in a business setting. The security, variety, and challenge are off the charts, and the breadth of techniques you'll encounter present a great environment for continuing to develop technically throughout your career while you grow as a professional. ■



New Version

StatXact 10

The Exact Statistical Software Standard

Celebrating 25 Years of Exact Statistics our 10th release adds a host of new statistical tests.

New in StatXact 10

- Genetic data: Family Based Association test
- Correlated Data: Exact tests for multinomial data.
- Boschloo's test, and power & sample size estimation
- Mid-p corrected confidence intervals
- Lepage test

Get more info at
www.cytel.com
 T+1.617.661.2011

Cytel