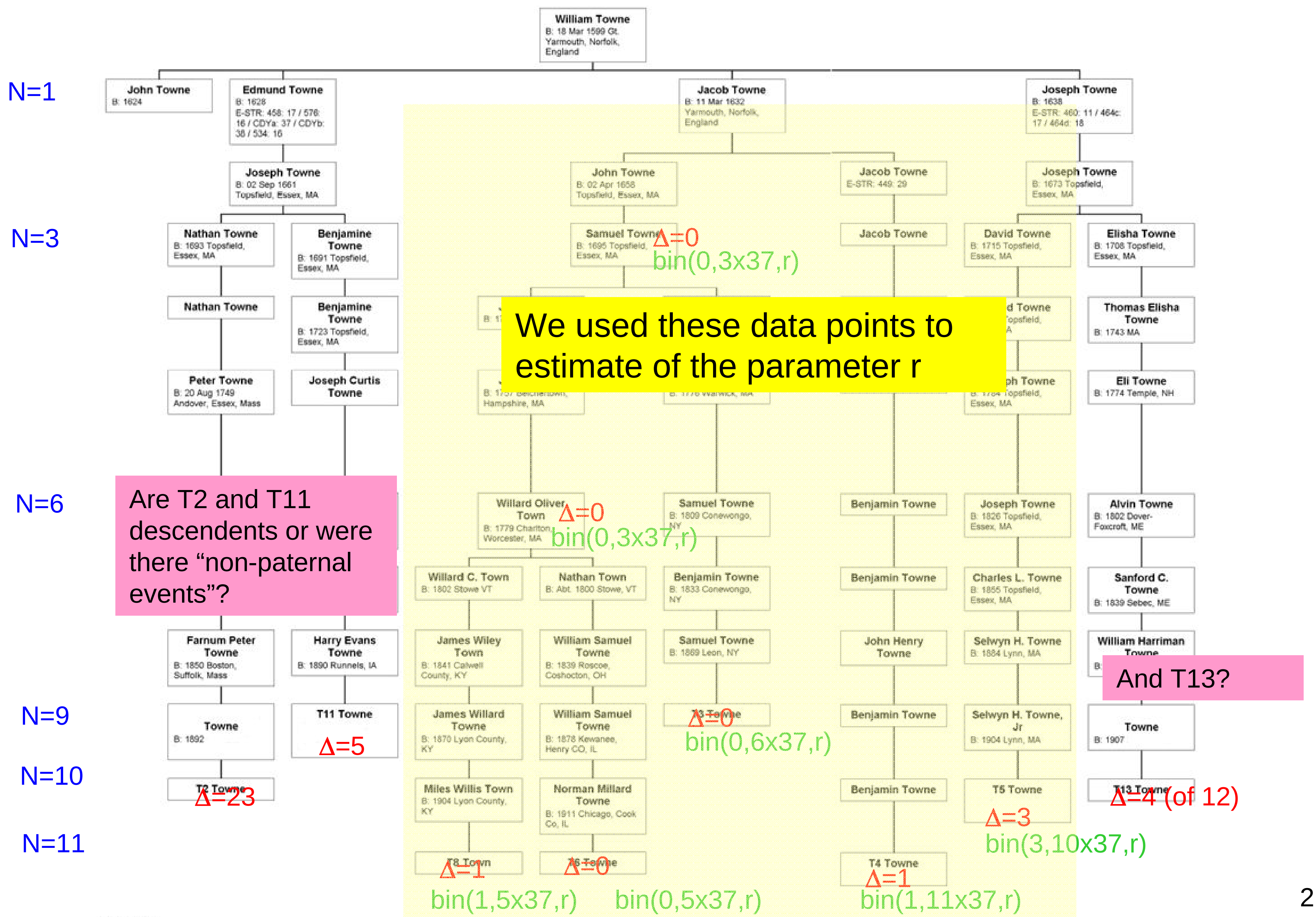
Opinionated Lessons in Statistics

by Bill Press

#15 The Towne Family – Again

Now that we're so adept at p-value stuff, let's go back to the Towne family.



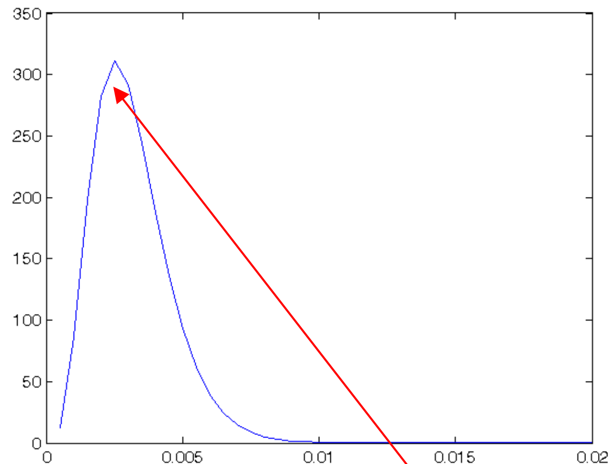
If we really knew *r*, then a p-value (tail) test on T2, T11, and T13 would be straightforward,

$$p_{\mathrm{tail},11} = \sum_{k=5}^{37} \mathrm{bin}(k, 9 \times 37, r)$$

notice how the "neglect backmutation" assumption comes in here

The problem is we have only Bayesian (uncertain) knowledge about *r*

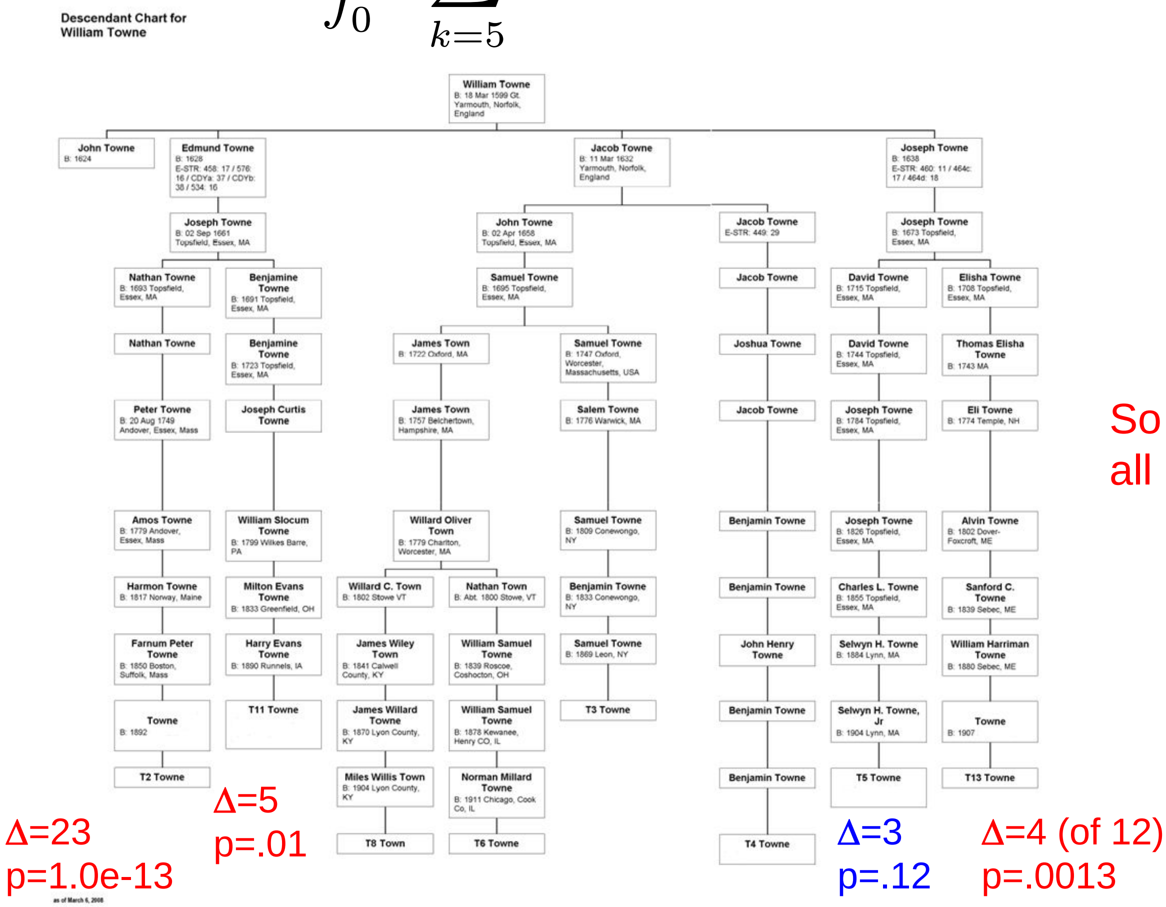
$$P(r|\mathrm{data}) = \mathrm{bin}(0, 3 \times 37, r)\,\mathrm{bin}(0, 3 \times 37, r)\,\mathrm{bin}(1, 5 \times 37, r)\,\mathrm{bin}(0, 5 \times 37, r)$$
$$\times\, \mathrm{bin}(0, 6 \times 37, r)\,\mathrm{bin}(1, 11 \times 37, r)\,\mathrm{bin}(3, 10 \times 37, r)/r$$



A common frequentist practice is to use the maximum likelihood estimate of r. **This is just wrong** (except asymptotically if the distribution of r were very narrow) because T11's extreme tail probabilities will be dominated by the extreme (but possible) values of r.

One “modern” way to proceed is to integrate the p-value over the posterior probability of all estimated quantities. This is called the “**posterior predictive p-value**” and is an example of a set of methods loosely called “**empirical Bayes**”.

$$p_{\mathrm{tail},11} = \int_0^\infty \sum_{k=5}^{37} \mathrm{bin}(k, 9\times 37, r)\, P(r|\mathrm{data})dr \Big/ \int_0^\infty P(r|\mathrm{data})dr$$



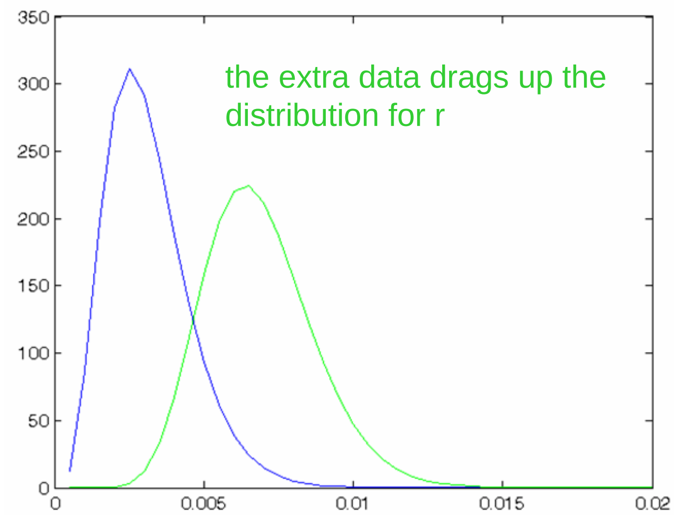
So the three questionables are all unlikely to be descendents.

This would be a satisfactory end to the Towne story, except that we tainted the data by tail trimming. While T2 is hopeless, what if we had included T11 and T13?

$$P(r|\text{data}) \propto P_{\text{previous}}(r|\text{data}) \times \text{bin}(5, 9 \times 37, r)\, \text{bin}(4, 10 \times 12, r)$$

```
t11tail =
    0.0953
t2tail =
  3.2348e-011
t13tail =
    0.0122
```

So suddenly there is hope for T11.
T2 and T13 still strongly ruled out.



This is an actual methodological problem with "posterior predictive p-value". Data is being used <u>twice</u>: once to get the posterior, then again to test itself. Often you can get away with this (e.g., try posterior both with and without questionable data). Doing so in this example, we find that T11 is left ambiguous.

This is when we need <u>real</u> 

(We'll return to the Towne family one more time, later.)

(Let me explain where we're going here…)

- Building up prerequisites to do a fairly sophisticated treatment of <u>model fitting</u>
  - Bayes parameter estimation ✓
  - p-value tail tests ✓
  - really understand multivariate normal and covariance
  - really understand chi-square
- Then, we get to appreciate the actual model fitting stuff
  - fitted parameters
  - their uncertainty expressed in several different ways
  - goodness-of-fit
- And it will in turn be a nice "platform" for learning some other things
  - bootstrap resampling