# MATHEMATICS AND COLLEGIATE WRESTLING:
## Wrestling with Byes

T. S. Michael
Mathematics Department
United States Naval Academy
Annapolis, MD 21402
tsm@usna.edu

National Joint Mathematics Meetings
Washington, DC
January 2009

NOTES FOR SLIDES

These notes support a talk presented in January 2009.
The corresponding slides are in a separate pdf file.

## Title Page
## Slide 0

I will discuss a real-world application of mathematics to an unusual scheduling problem in college wrestling.

---

## Summary
## Slide 1

A major college wrestling conference faces a problem each year.
How do they allocate the byes in the first round of the conference tournament?
The old method sometimes leads to unsatisfactory, 'unbalanced' situations.

I proposed a new method based on using a 'backwards least squares' calculation.
This method has been implemented since 2006.

For more information and wrestling background, email me.
Slides, notes, and answers to frequently asked questions will appear on my website.

---

## EIWA
## Slide 2

The EIWA is the oldest and one of the best wrestling conferences in the country. There are 13 teams and 10 weight classes.

Technical issue: The conference recently went up to 14 teams. Our model has been adapted to that situation, too.

The EIWA tournament is a qualifying meet for the NCAA meet.
Only individual wrestlers can qualify for nationals, not teams.
Each team enters (at most) one wrestler in each weight class.

Wrestling is an individual sport, but at the end of the tournament the EIWA team champion is crowned based on scoring in each weight class.

Each weight class is its own separate tournament.
A standard championship bracket is used.
There will be three byes in the first round since $13 = 16 - 3$.

## Question
## Slide 3

How do we decide who gets the three byes in the first round?

The answer might surprise people not familiar with wrestling.

The 8 best wrestlers in each weight class are ranked (seeded) 1–8 before the tournament starts.
In each weight class the 3 byes are distributed at random among the 8 seeded wrestlers.

Each seed has a 3-out-of-8 chance of getting a bye.
So the bye-probability is 0.375 for each seed.

---

## How to Make an EIWA Wrestling Bracket
## Slides 4–6

An EIWA wrestling bracket is constructed in 3 steps.

First: Place the 8 seeds in the expected positions in the bracket.
Seeds 1 and 2 will meet in the finals if there are no upsets.

Second: Distribute the 3 first-round byes randomly among the 8 seeds.

Technical issue: However, the 3 byes are forbidden to all occur in the same half of the bracket; the two halves of the bracket must be 'balanced" with respect to byes.

Third: Pair the 5 remaining seeds with unseeded wrestlers randomly.

## 2002: Seeds and Byes
## Slide 7

This table shows the seeds and byes for the 2002 tournament.
Bucknell had 2 seeds, and therefore expected 0.75 byes
since each seed had a 3-out-of-8 chance of getting a bye.

The top teams had 9 or 10 seeds and expected to get about 4 byes.
The number of byes received by the teams was about the number they expected.

Notice that there were 35 byes altogether, rather than 30.
This is because some teams did not enter wrestlers in every weight classes.
This situation occurs most years.

Most years things go smoothly.
But the independent and random placement of byes can lead to surprising events....

## 2004: Seeds and Byes
## Slide 8

Every few years the random draw is unbalanced for the teams.
This table shows what happened in 2004.
The top three teams expected about 4 byes each.
But Cornell got 7 byes.
And Penn got 0 byes.

This is a problem for the team championship.
Wrestlers can score 'bonus points' for their team by pinning their opponents, say.
A team with an inordinate number of byes has too few chances to get those first-round bonus points.
It would be better to have more balanced draws.

The standard combinatorial matching and optimization algorithms do not work!
Reason: They skew the bye-probabilities for seeded wrestlers away from 0.375.

## Graphic Display of Data
## Slide 9

Let's display the data graphically.
The top graph illustrates the 2002 data.
The horizontal axis is for byes expected.
The vertical axis for byes actually received.
Each red dot describes a team.
Most dots are clustered near the blue 45-degree line in 2002.
This reflects the goodness of the draw in 2002.

The bottom graph shows 2004.
We can see visually that the draw was so poor.
Several red dots are far away from the ideal line.

These graphs remind us of the familiar least squares procedure.
But our situation is backwards.
Instead of trying to find the best line to fit the data points,
we prefer to have the randomly generated points fall near the ideal line.

## The Penalty of a Draw
## Slide 10

We can measure how good or bad a draw is with a single number.

Suppose we have a complete draw—brackets for all 10 weight classes.
Each team has a discrepancy—the difference between the number of byes it expected and the number it actually received.

The penalty is the sum of the squares of the discrepancies over all teams.

A large penalty means a poor distribution of byes among teams.
A small penalty means a good distribution of byes among teams.

## New Method (2006-present)
## Slide 11

Since 2006 EIWA has used a method I proposed to distribute byes.
Here's how it works.

First: Make a complete random draw.
This is just like the old method.

Second: Compute the penalty of the draw.

Third: If the penalty exceeds a pre-determined cut-off,
then reject the draw. Start over.

Fourth: If the penalty does not exceed a pre-determined cut-off,
then accept the draw. Start wrestling!

## Penalty Distributions
## Slide 12

The table shows the penalties for four recent tournaments.
The bar graph shows results of a simulation.
10,000 random draws were made using seed data from 2004.
The penalties are distributed are distributed as shown.
The median penalty is about 18.
The mean penalty is about 19.
The distribution is similar with seed data from other years.

## 2006: First Draw
## Slide 13

Here's what happened in 2006.
The table shows the first draw.
The penalty is large, about 39—much larger than the cut-off.
With the old method, this unbalanced draw would be used.

With the new method, this draw is rejected.....

## 2006: Fifth Draw
## Slide 14

Here is the fifth and final draw.
The penalty is small, about 9—much smaller than the cut-off.
This was the draw used for the tournament.

## The Penalty Cut-Off (Extreme Cases)
## Slide 15

Let's talk about how the cut-off is determined in advance.

Let's look at extreme cases first.
If the cut-off is $\infty$, then we accept the first draw.
This is the old method.

If we set the cut-off close to $0$,
then we are using a 'quota method'.
We are demanding that the number of byes for each team
is (nearly) equal to the number of expected byes.
This is a reasonable mathematical urge.
But this method skews the bye-probabilities
of the seeded wrestlers away from 0.375.
This is unacceptable.

For example, consider the lone seeded Princeton wrestler from 2006.
Assigning him $0$ byes, means the $8$ seeds in his weight class are treated differently.
He has no chance of a bye, while the other $7$ seeds have 3-out-of-7 chance for a bye.

We aren't allowed to treat the 8 seeds differently in a weight class
since they are all competing for the same berths at the national meet.

## How to Choose the Penalty Cut-Off Number
## Slide 16

We use an intermediate value of the cut-off number.
This greatly reduces unfairness to teams,
while preserving fairness to individual wrestlers.

The cut-off is chosen so that each seeded wrestler's bye probability
is skewed by less than 0.01.

Extensive simulations were performed for various scenarios.
The number of teams could be 13 or 14.
The number of byes could range from 20 to 36.
For each scenario a cut-off is found that skews the bye-probabilities $< 0.01$.