

# Mendelian Sampling Variance Software Version 3.0 - new features and improvements

Anna-Maria Leino, Martin Lidauer, and Esa Mäntysaari\*

Animal Genetics,  
Production Systems,  
Luke

# Long road from a need and an idea to IB4

- Experience has shown that genetic variance trends in national evaluations affect MACE evaluations:
  - Top bulls from birth year classes with inflated genetic variances benefit from an upward bias in ranking
  - Validation of the national evaluation models for the heterogeneity of genetic variance needed
- How to obtain yearly genetic variance?
- Sullivan (1999) suggested a REML method for estimating within-year genetic variances by deriving an equation based on Mendelian sampling and its prediction error variance

# Long road from a need and an idea to IB4

- 2003: Beltsville meeting, Mendelian sampling method for validation
- 2003, 2005: First version of the validation test presented by Fikse et al., Interbull Bulletin
- 2007: Lidauer et al. suggested a FMS method
- 2011: New research project to continue work with MS validation
  - Tyrisevä et al. 2011, 2012. Interbull Bulletin.
  - Software for Mendelian sampling variance test
- 2014: First pilot test of Interbull member countries
- 2015-2016: MS working group active, analyses of the pilot test results, some modifications to the software, Version 2.6
- 2017: The MS variance validation test was accepted for official use
  - During first years the test results have no consequences for countries, but results are collected to learn more about the behavior of the test

# Documentation of the method



**J. Dairy Sci. 101:2187–2198**  
**<https://doi.org/10.3168/jds.2017-13255>**

© American Dairy Science Association®, 2018.

## Validation of consistency of Mendelian sampling variance

**A.-M. Tyrisevä,\*<sup>1</sup> W. F. Fikse,†<sup>2</sup> E. A. Mäntysaari,\* J. Jakobsen,‡ G. P. Aamand,§ J. Dürr,# and M. H. Lidauer\***

\*Natural Resources Institute Finland, Green Technology, Biometrical Genetics, 31600 Jokioinen, Finland

†Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, 75007 Uppsala, Sweden

‡Norwegian Association of Sheep and Goat Breeders, 1431 Ås, Norway



**J. Dairy Sci. 101:3155–3163**  
**<https://doi.org/10.3168/jds.2017-13527>**

© American Dairy Science Association®, 2018.

## Detection of evaluation bias caused by genomic preselection

**A.-M. Tyrisevä,\*<sup>1</sup> E. A. Mäntysaari,\* J. Jakobsen,† G. P. Aamand,‡ J. Dürr,§ W. F. Fikse,#<sup>2</sup> and M. H. Lidauer\***

\*Natural Resources Institute Finland (Luke), Green Technology, Biometrical Genetics, 31600 Jokioinen, Finland

†Norwegian Association of Sheep and Goat Breeders, 1431 Ås, Norway

‡NAV Nordic Cattle Genetic Evaluation, 8200 Aarhus, Denmark

§Council on Dairy Cattle Breeding, Bowie, MD 20716

#Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, 75007 Uppsala, Sweden

## The original method shortly

$$\sigma_{u_i}^2 = \frac{\sum_{k=1}^{q_i} d_k \hat{m} s_k^2}{q_i - \sum_{k=1}^{q_i} d_k PEV(\hat{m} s_k)}$$

- Estimates within-year genetic variances using MS term and its prediction error variance
- Fits a weighted linear regression between the estimates and the years under study
- **Identify possible outliers that do not fit the linear trend model**
- Bootstrap a 95% empirical confidence interval for a possible trend in the estimate
- Fit tolerance thresholds both for the trend and the outliers to detect only cases that have real practical impact:
  - Trend:  $\pm 2\%$
  - Outliers:  $\pm 10\%$   
of the average estimated variance

# Mendelian sampling variance test

## Version 3.0

- ITC recommended improvements after studying the validation test results of September 2017 test run:
  1. "Include an average value for MS mean in the same fashion as it is done for the genetic variance and include it in the summary output"
    - There is an interest whether the test can be utilized to detect genomic pre-selection
  2. "Modify the outlier test so that the outlier will be detected from zero trend rather than from the trend line"

## Version 3.0: new output

TRAIT: t3

Year	N Animals	Estimated Genetic Variance	Mendelian Sampling Mean	MS mean / sqrt(G)	Outlier test: Lower Limit	Upper Limit
1988	259	139.32	-1.8741	-0.15	105.55	172.50
1989	241	130.75	-0.91577	-0.07	95.398	175.30
1990	216	115.78	0.18495	0.01	85.895	149.73
1991	273	146.92	-1.1108	-0.09	109.74	191.75
1992	275	120.09	-0.77382	-0.06	89.253	156.75
1993	237	138.35	-1.4272	-0.11	97.796	180.03
1994	278	160.23	-2.2446	-0.18	112.31	208.09
1995	285	165.30	-2.5911	-0.21	122.70	218.60
1996	333	142.87	-1.6991	-0.14	104.37	188.29
1997	345	217.69	-3.1251	-0.25	162.05	289.21
1998	295	185.86	-2.6683	-0.21	144.94	237.71
1999	180	206.93	-3.4308	-0.27	147.53	267.18

FAILED

FAILED

Total number of the animals in the analysis: 3217

Average genetic variance: 155.67

Average MS mean: -1.8063

## ESTIMATED TREND IN THE GENETIC VARIANCE

Regression coefficients (B0, B1) of the weighted linear regression on year.

B0: 115.99 B1: 7.2651 Yearly trend (% relative to genetic variance):

Empirical 95% confidence interval:

Lower limit: 3.1 Upper limit: 6.3

Trend deviated statistically significantly from zero and was also above the tolerated threshold of +-2% ( 4.7 )

## OUTLIER TEST:

95% confidence intervals for the yearly genetic variance estimates.

The confidence intervals should include the average genetic variance: 155.67

Outliers exceeding the statistical and tolerated thresholds are marked with "failed".

Yearly MS means expressed in units of average genetic standard deviation

Average MS mean across years added

One table instead of two separate

# Version 3.0: new outlier test

TRAIT: t3

Year	N Animals	Estimated Genetic Variance	Mendelian Sampling Mean	MS mean / sqrt(G)	Outlier test: Lower Limit	Upper Limit	
1988	259	139.32	-1.8741	-0.15	105.55	172.50	
1989	241	130.75	-0.91577	-0.07	95.398	175.30	
1990	216	115.78	0.18495	0.01	85.895	149.73	FAILED
1991	273	146.92	-1.1108	-0.09	109.74	191.75	
1992	275	120.09	-0.77382	-0.06	89.253	156.75	
1993	237	138.35	-1.4272	-0.11	97.796	180.03	
1994	278	160.23	-2.2446	-0.18	112.31	208.09	
1995	285	165.30	-2.5911	-0.21	122.70	218.60	
1996	333	142.87	-1.6991	-0.14	104.37	188.29	
1997	345	217.69	-3.1251	-0.25	162.05	289.21	FAILED
1998	295	185.86	-2.6683	-0.21	144.94	237.71	
1999	180	206.93	-3.4308	-0.27	147.53	267.18	

Total number of the animals in the analysis: 3217

Average genetic variance: 155.67

Average MS mean: -1.8063

## ESTIMATED TREND IN THE GENETIC VARIANCE

Regression coefficients (B0, B1) of the weighted linear regression on year.

B0: 115.99 B1: 7.2651 Yearly trend (% relative to genetic variance): 4.67

Empirical 95% confidence interval:

Lower limit: 3.1 Upper limit: 6.3

Trend deviated statistically significantly from zero  
and was also above the tolerated threshold of +-2% ( 4.7 )

## OUTLIER TEST:

95% confidence intervals for the yearly genetic variance estimates.

The confidence intervals should include the average genetic variance: 155.67

Outliers exceeding the statistical and tolerated thresholds are marked with "failed".



# Version 3.0: new outlier test

- An empirical 95% confidence interval is obtained for each yearly genetic variance estimate by bootstrapping 1000 samples with replacement and calculating within-year genetic variance estimates for each bootstrap sample
- If the empirical CI for tested year  $i$  does not include **the average genetic variance**, the year is labeled as statistical outlier.
- Also tolerance thresholds are considered to label only cases that have a practical importance:  $\sigma_{u_i}^2 \pm 0.10\bar{\sigma}_u^2$

## OUTLIER TEST:

95% confidence intervals for the yearly genetic variance estimates.  
The confidence intervals should include the average genetic variance: 155.67  
Outliers exceeding the statistical and tolerated thresholds are marked with "failed".

# Examples of the results obtained under versions 2.6 and 3.0

Explanations of the graphs that can be printed by using R and/or SAS codes provided with the package

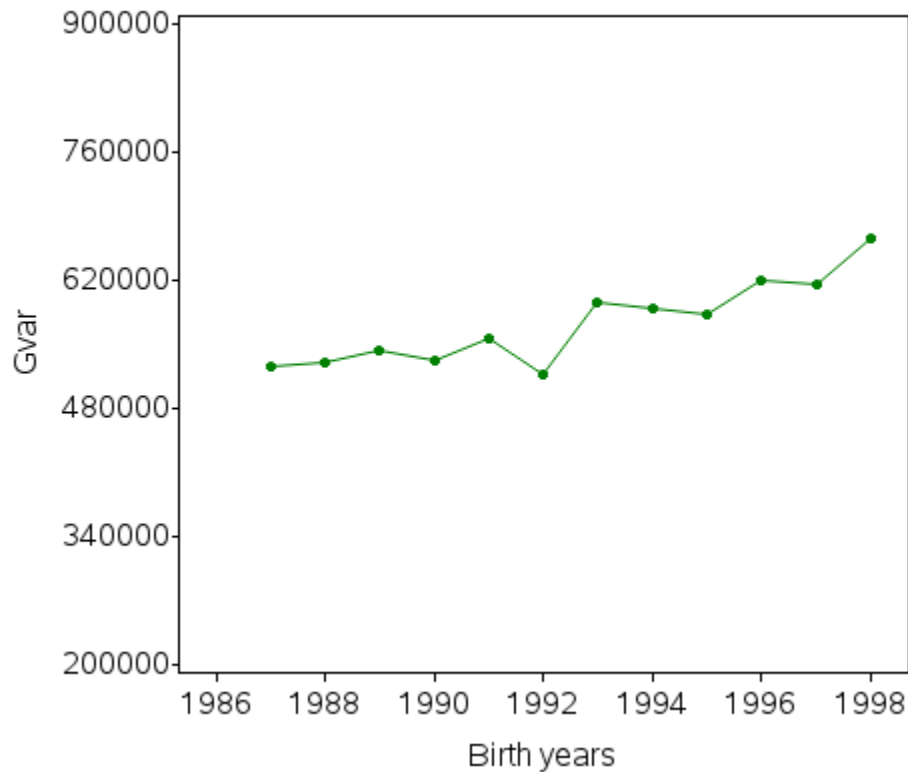
- If the trend in genetic variance is statistically non-significant or significant but within the tolerated threshold of  $\pm 2\%$  **trend line is green,**
- Otherwise (potential problem detected) **trend line is red**
- Statistical outliers that are also outside the tolerance thresholds are marked with **out**

Graph improvements from Version 2.6:

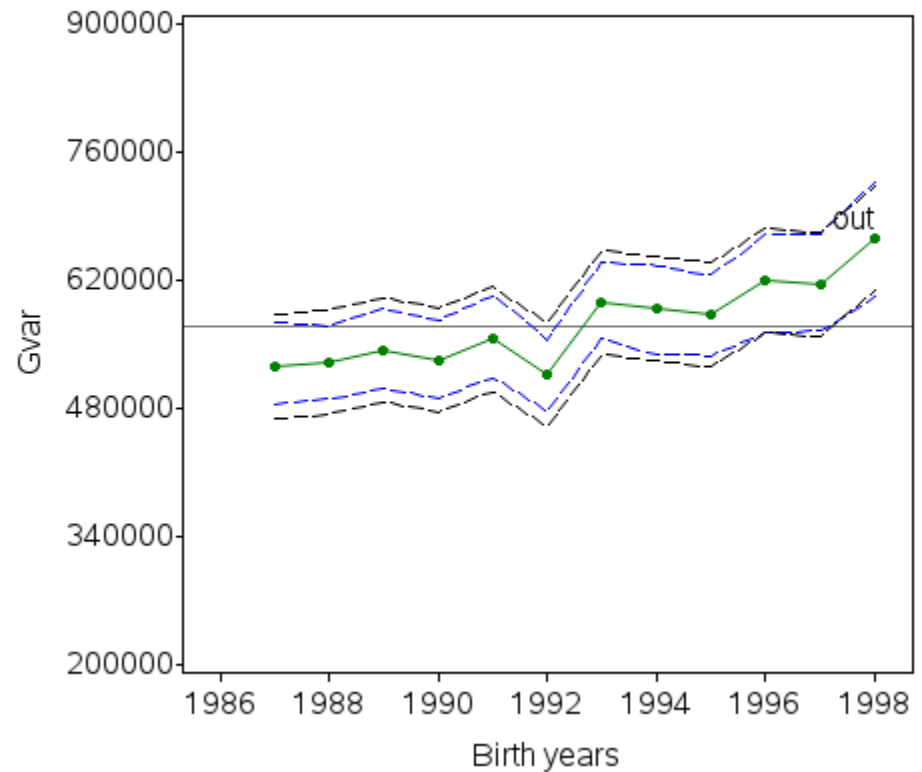
- Average genetic variance line added
- 95% CI for outliers added, color is blue
- Tolerance thresholds for outliers added, color is black

# Examples of the results obtained under versions 2.6 (left hand side) and 3.0 (right hand side)

Within-year genetic variances



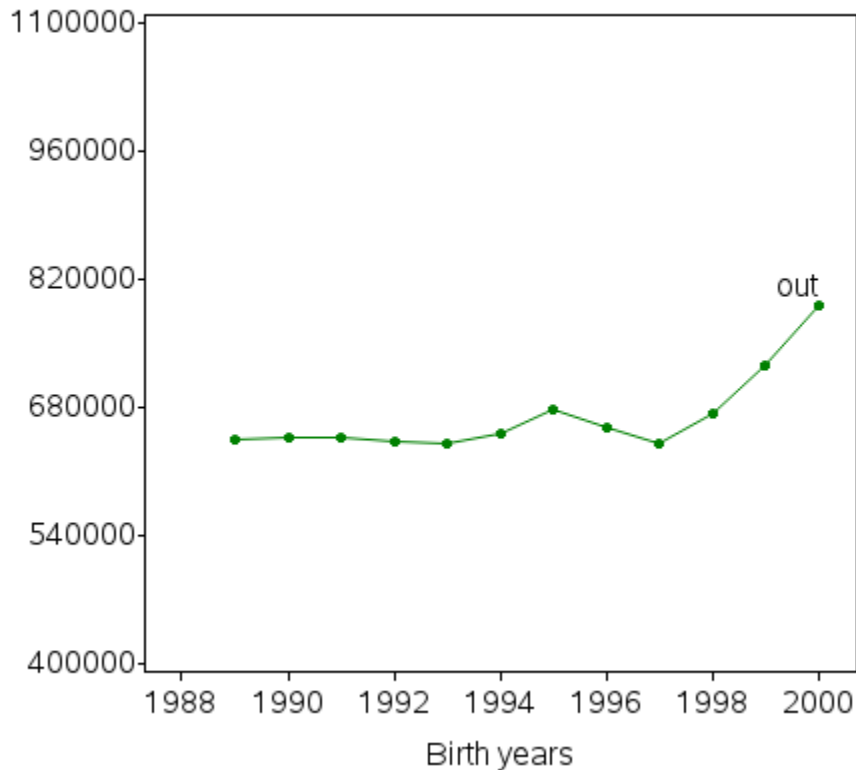
Within-year genetic variances



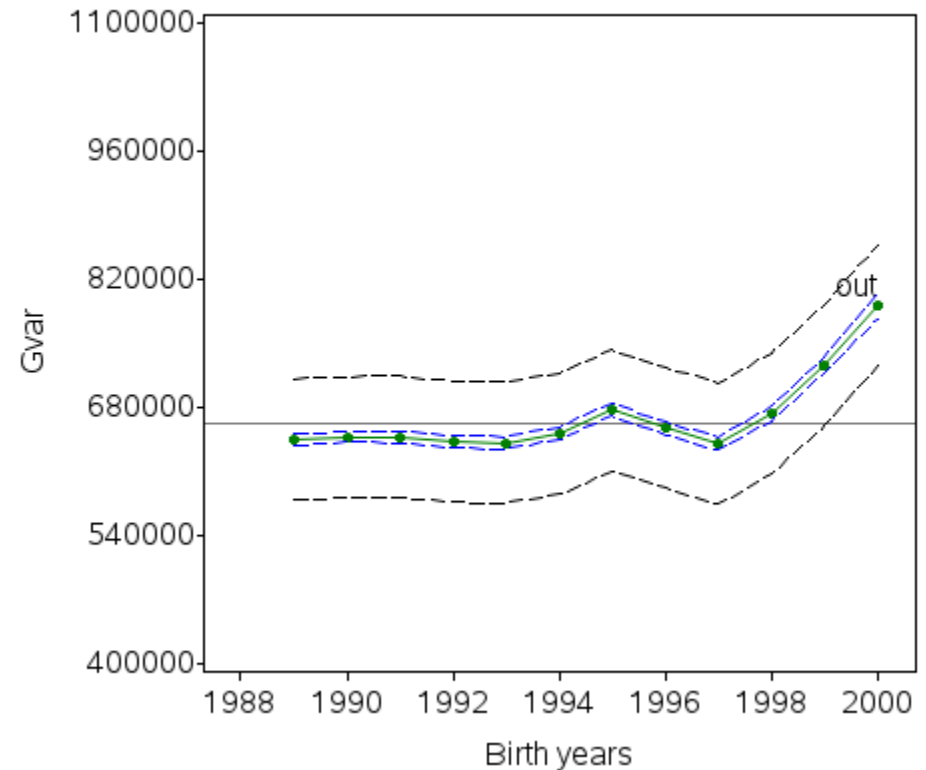
Data: bulls

# Examples of the results obtained under versions 2.6 (left hand side) and 3.0 (right hand side)

Within-year genetic variances



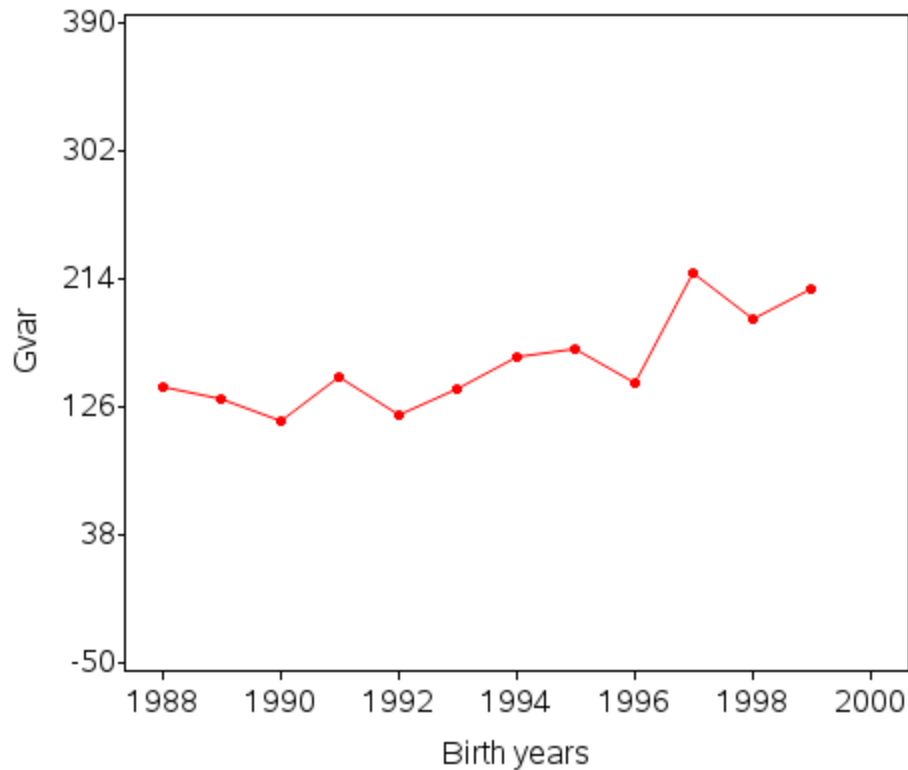
Within-year genetic variances



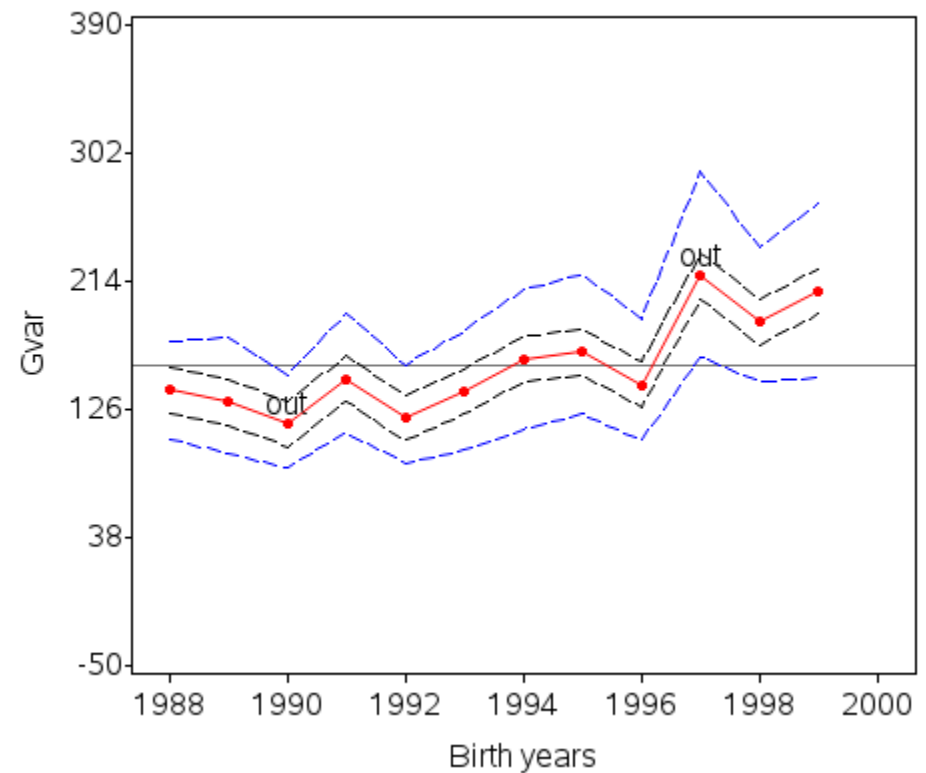
Data: cows

# Examples of the results obtained under versions 2.6 (left hand side) and 3.0 (right hand side)

Within-year genetic variances



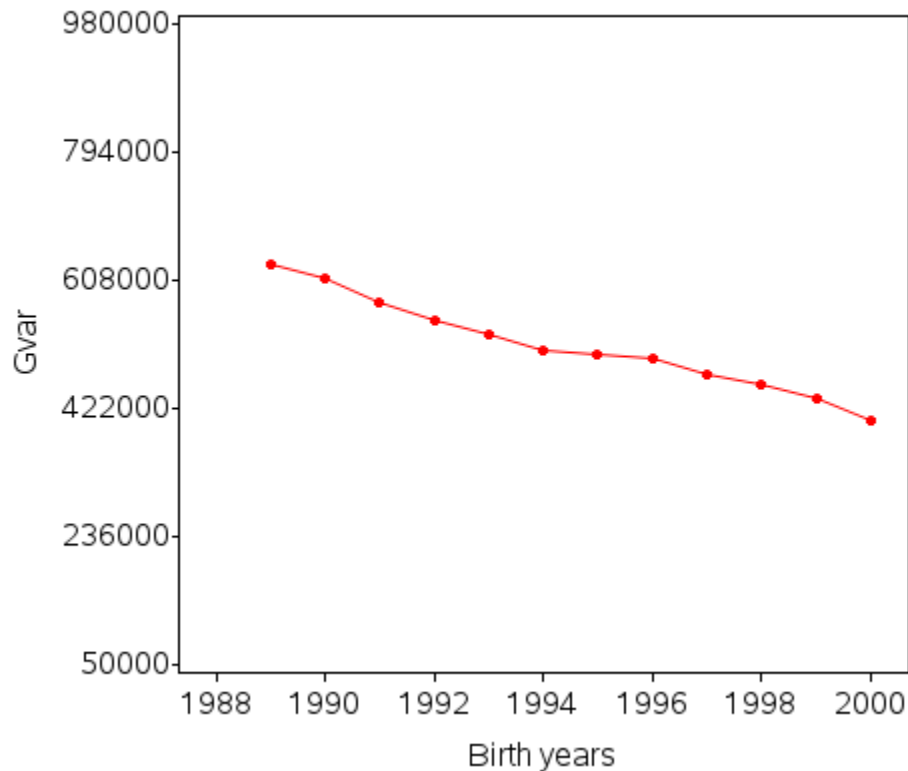
Within-year genetic variances



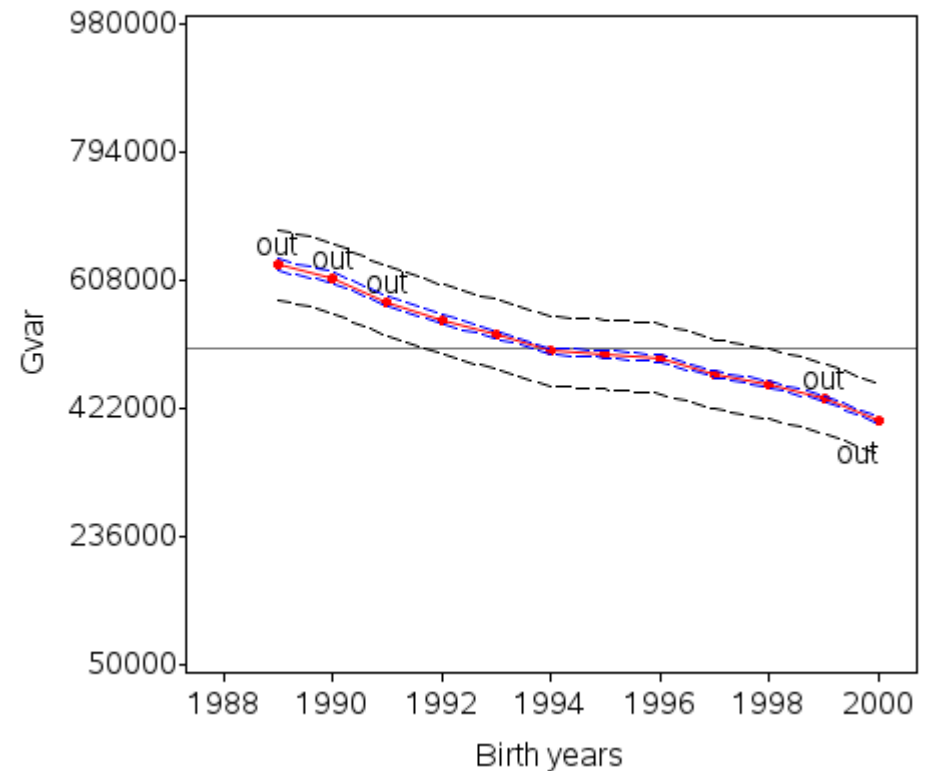
Data: bulls

# Examples of the results obtained under versions 2.6 (left hand side) and 3.0 (right hand side)

Within-year genetic variances



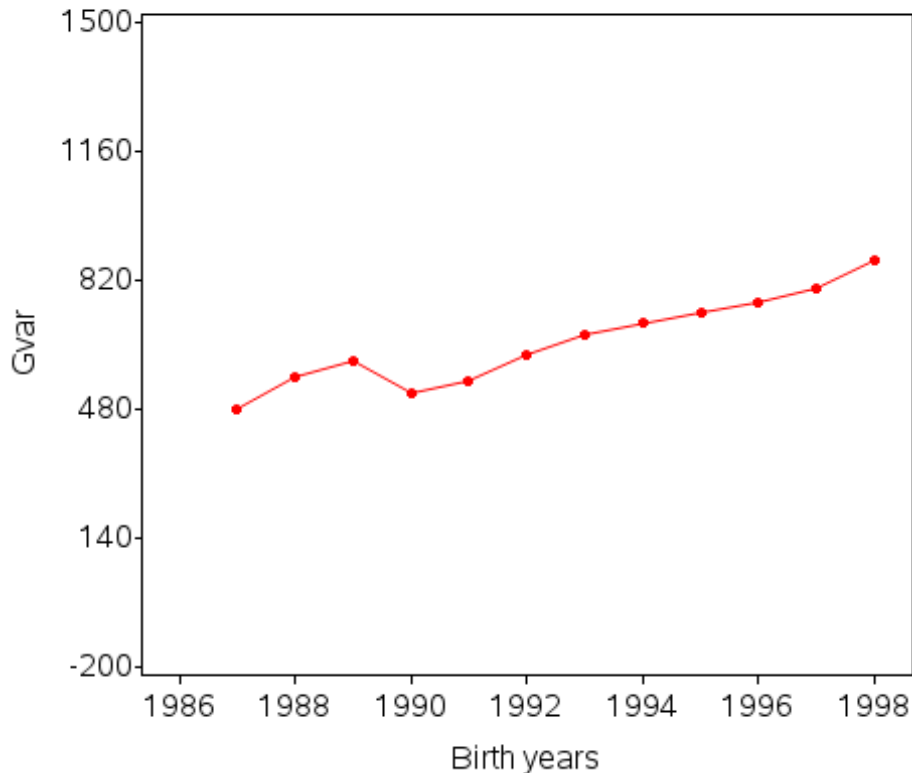
Within-year genetic variances



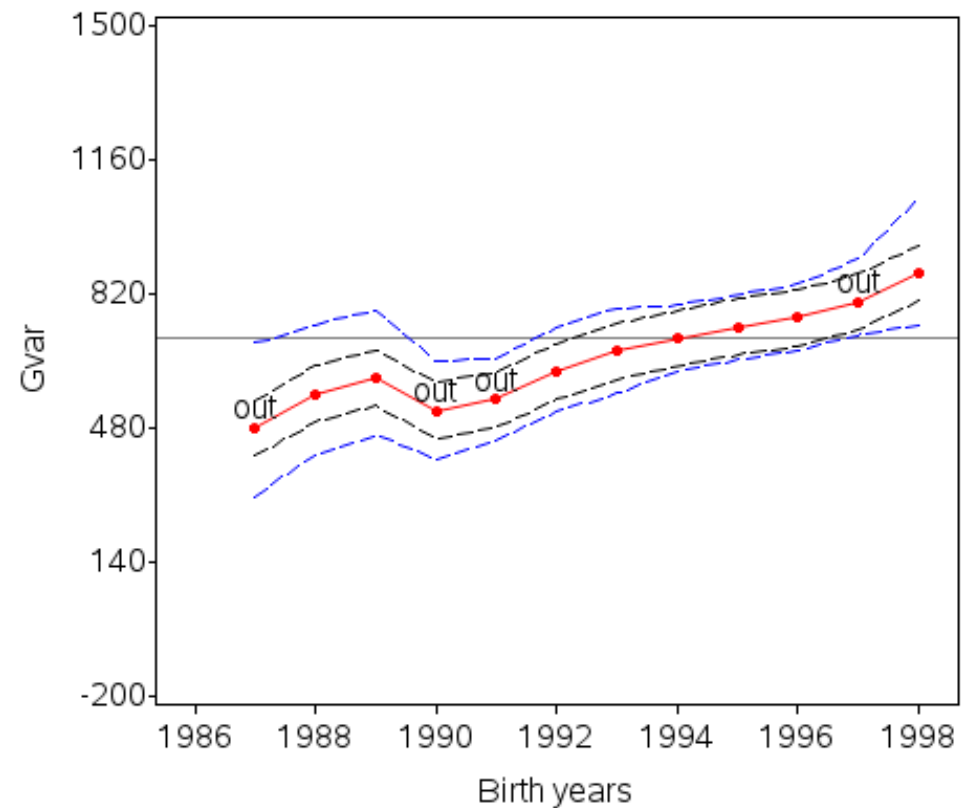
Data: cows

# Examples of the results obtained under versions 2.6 (left hand side) and 3.0 (right hand side)

Within-year genetic variances



Within-year genetic variances



Data: bulls

Thank you!