Abstracts

2017 ASAS—CSAS Annual Meeting and Trade Show

July 8–12, 2017
Baltimore, MD

American Society of Animal Science
Journal of Animal Science
Volume 95, Supplement 4

When the misclassified data was analyzed with M1, a 20.5% and 11.8% bias was observed in the estimation of the heritability for the large and small data sets, respectively. Using M2, bias was removed. In fact, estimates of heritability were almost identical to those obtained using the real data (0.106 vs. 0.106 and 0.097 vs 0.098 for 10K and 1.5K data sets, respectively). Furthermore, the proposed method was able to detect true misclassified records with high probability. These results clearly indicate the effectiveness of the proposed method in reducing bias in the analysis of discrete data subject to misclassification.

Key Words: calving ease, misclassification, multinomial responses
doi:10.2527/asasann.2017.187

188 Ability to genotype differing variants with arrays vs. whole genome sequencing. P. M. VanRaden*[1], G. L. Spangler[1], C. P. VanTassell[1], J. Jiang[2], L. Ma[2], J. R. O'Connell[3], S. Smith[4], and S. K. DeNise[4], *[1]Animal Genomics and Improvement Laboratory, ARS, USDA, Beltsville, MD, [2]University of Maryland, College Park, [3]University of Maryland-Baltimore, Baltimore, [4]Zoetis Inc., Kalamazoo, MI.*

Whole genome sequencing has identified millions of new variants, but many (about 35% in our experience) of the single nucleotide polymorphisms (SNPs) may not produce high quality genotypes from microarrays. Properties of SNPs can help predict which will pass or fail when designing arrays, such as the customized version of Illumina's Bovine LD chip examined here. Genotypes for 26,970 reference bulls were imputed using 440 sequenced Holsteins from run 5 of the 1000 Bull Genomes Project, and 4,821 SNPs with largest effects for net merit were selected. When adding those to the Zoetis LD chip (version 5), the success rate was 96% for 3,220 SNPs from the Bovine HD chip, but only 64% for 1,601 new sequence SNPs not previously on any chip. To determine why SNPs failed, a pass/fail (1/0) indicator of sequence SNP conversion success was correlated with (1) Illumina design scores, (2) estimated heritabilities of the genotypes for 3,000 randomly selected bulls, and (3) the base distance that the SNP was inside a repetitive DNA segment as determined by RepeatMasker, using a minimum distance of 0 if outside a repeat and maximum of 50 bases if inside. The correlations were 0.51 for design scores, 0.14 for estimated heritabilities, and −0.15 for repeat distance. All three were highly significant ($P < 0.0001$), but repeat distance was less significant ($P = 0.04$) after fitting design score and heritability in multiple regression. Three other factors (minor allele frequency, SNP position with genes, and the reference/alternate allele combination pattern) were not associated with conversion success. In a reverse test, 56,815 SNPs from the Bovine 50K version 1 chip were matched with 38 million sequence SNPs. Previously 15,772 of the 50K SNPs had been declared not usable, and 11,969 (87%) of those were also either not identified or removed by sequence edits. However, 3,803 (9%) of the 43,053 currently used SNPs that produce high quality genotypes on the 50K chip were absent from the sequence data, and the absence was not associated with minor allele frequency or allele combination. If the goal is to select the best SNP subset for a chip, design scores could be pre-computed and examined before rather than after estimating SNP effects, allowing selection of other linked SNPs expected to perform better. Eventually, targeted sequencing could provide genotypes for important SNPs that fail to convert because many SNPs from sequence data are difficult to genotype using arrays.

Key Words: design scores, genotyping arrays, single nucleotide polymorphisms
doi:10.2527/asasann.2017.188

189 Genomic relatedness strengthens genetic connectedness across management units. H. Yu*, M. L. Spangler, R. M. Lewis, and G. Morota, *University of Nebraska-Lincoln, Lincoln.*

Genetic connectedness refers to a measure of genetic relatedness across management units (e.g., herds and flocks) in animal breeding. Connectedness has shown to be an important measure of reliability when comparing genetic values derived from pedigree-based best linear unbiased prediction (BLUP) among management units. With the presence of high genetic connectedness in management units, BLUP is known to provide less biased comparisons between genetic values. Genetic connectedness has been applied successfully to pedigree-based BLUP; however, relatively little attention has been paid to using genomic information, such as single nucleotide polymorphisms, to estimate genetic connectedness. Thus, it remains unclear whether and to what extent genome-based information enhances connectedness. In this study, we assessed genome-based genetic connectedness across management units by applying prediction error variance of difference (PEVD), coefficient of determination (CD), and prediction error correlation (r) to a combination of computer simulation and real data (mice and cattle). Relationship matrices were constructed from three different sources: pedigree (A), genomics (G), and a hybrid of these two. We found that genomic information increased the estimate of connectedness among individuals from different management units compared to that of pedigree, and a disconnected design benefited the greatest. In the well-structured mice data (full-sib families), regardless of heritability (0.2 vs. 0.8), all 3 statistics inferred increased connectedness across-units when using G- rather than A-based relationships. With the cattle data, genomic relationships decreased PEVD across-units suggesting stronger connectedness. With r once scaling G to values between 0 and 2, which is intrinsic to A, connectedness also increased with genomic information. However, PEVD often increased and r often decreased when obtained using the alternative form of G, instead suggesting less connectedness. Such inconsistencies were not found with CD. Caution should be exercised