

Table S1. Comparison of the variability of empirical ML estimates $\tilde{\theta}$ from empirical data sampled with different n values. Instead of deriving $\hat{\theta}$ from a single (almond) dataset, we iteratively sample (10^4 times) $\tilde{p}$ from a product multinomial with input $\hat{p}$ (i.e. treated as known population proportions) and tested chosen n value, and per each iteration we derive $\tilde{\theta}$ ML estimates of $\hat{\theta}$ and corresponding Wald confidence intervals of $\hat{\theta}$ from the Hessian matrix.

	True sample size n			
	50	150	250	500
mean($\tilde{a}_1$)	695.54073	695.65269	695.60218	695.58277
$\pm$ mean($z_{0.025}\sqrt{\tilde{v}_{11}}$)	± 8.17112	± 4.72359	± 3.65994	± 2.58823
$\tilde{a}_1$ empirical coverage of 95% CI	94.84	95.00	95.16	95.18
mean($\tilde{a}_2$)	769.74950	769.80461	769.75212	769.73440
$\pm$ mean($z_{0.025}\sqrt{\tilde{v}_{22}}$)	± 8.61470	± 4.97883	± 3.85836	± 2.72896
$\tilde{a}_2$ empirical coverage of 95% CI	94.70	94.73	94.89	94.77
mean($\tilde{a}_3$)	816.26883	816.25789	816.23719	816.22629
$\pm$ mean($z_{0.025}\sqrt{\tilde{v}_{33}}$)	± 7.82495	± 4.52385	± 3.50579	± 2.47949
$\tilde{a}_3$ empirical coverage of 95% CI	95.02	95.01	94.96	94.66
mean($\tilde{a}_4$)	919.19294	919.11557	919.03054	919.06322
$\pm$ mean($z_{0.025}\sqrt{\tilde{v}_{44}}$)	± 10.22141	± 5.90681	± 4.57643	± 3.23628
$\tilde{a}_4$ empirical coverage of 95% CI	94.84	95.10	93.98	94.99
mean($\tilde{a}_{r-1}$)	953.03190	953.03091	952.97243	952.96517
$\pm$ mean($z_{0.025}\sqrt{\tilde{v}_{r-1r-1}}$)	± 9.42821	± 5.44798	± 4.22180	± 2.98584
$\tilde{a}_{r-1}$ empirical coverage of 95% CI	94.87	94.94	94.87	94.78
mean($\tilde{v}$)	1.07785	1.08263	1.08393	1.08462
$\pm$ mean($z_{0.025}\sqrt{\tilde{v}_{rr}}$)	± 0.18979	± 0.10988	± 0.08517	± 0.06025
$\tilde{v}$ empirical coverage of 95% CI	94.54	94.56	94.29	95.08