STAT 457 Homework 05

Martha Eichlersmith

2019-12-03

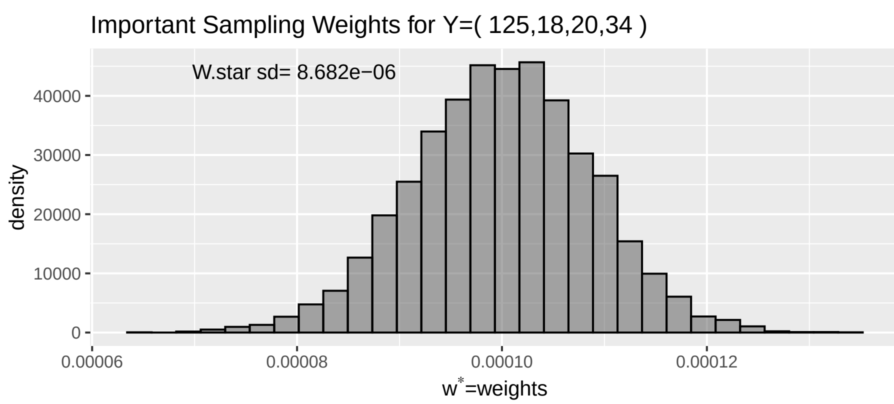
Problem 3

For the genetic linkage model:

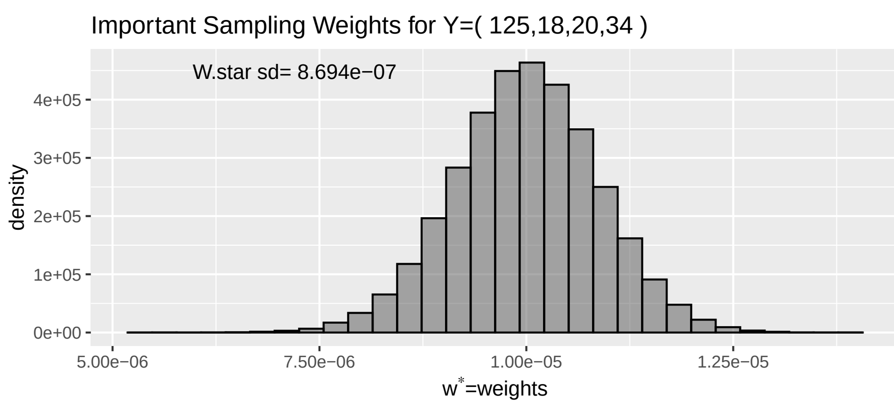
Problem 3a

Use importance sampling to obtain the posterior mean for data $Y = (125, 18, 20, 34)$. Use the matching normal distribution as the importance function. Compare your importance sampling estimates of the posterior mean to those obtained via Laplace's method. Draw the histogram of the weights and compute their standard deviation.

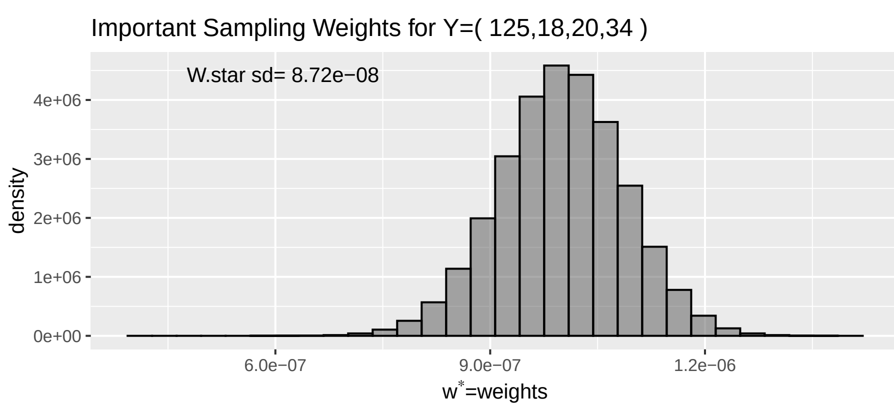
$$\text{Normal Approximation for } Y = (125, 18, 20, 34) \sim \mathcal{N}(\mu = 0.62682,\ \sigma = 0.05382)$$



	IS	Norm.Apprx	Diff
mean	0.62287	0.62682	−0.00395
sd	0.05325	0.05382	−0.00057



	IS	Norm.Apprx	Diff
mean	0.62173	0.62682	−0.00509
sd	0.05349	0.05382	−0.00033

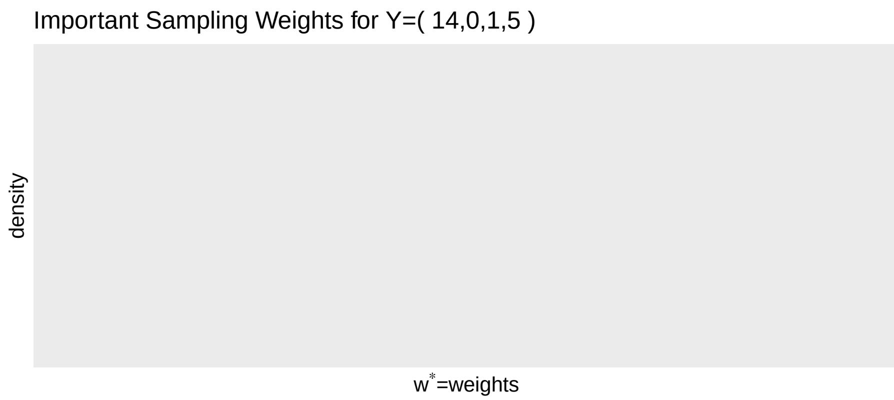


	IS	Norm.Apprx	Diff
mean	0.62211	0.62682	−0.00471
sd	0.05356	0.05382	−0.00026

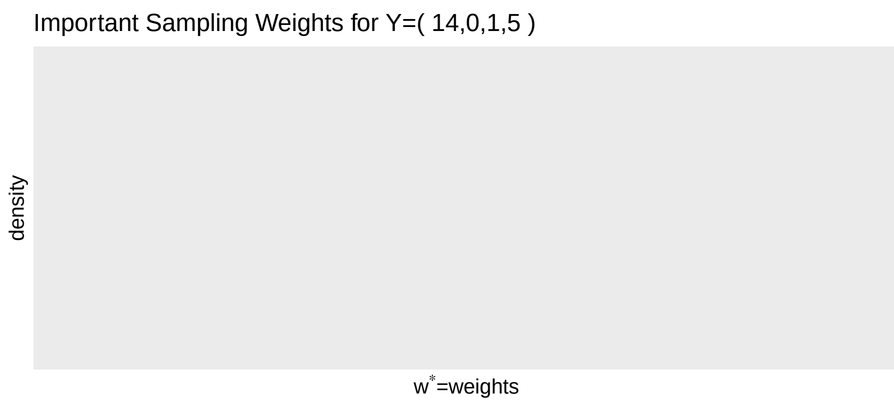
Problem 3b

Repeat (a) for the data $Y = (14, 0, 1, 5)$. Compare the histograms of the weights for both data sets. Discuss the adequacy of important sampling estimate for each data set.

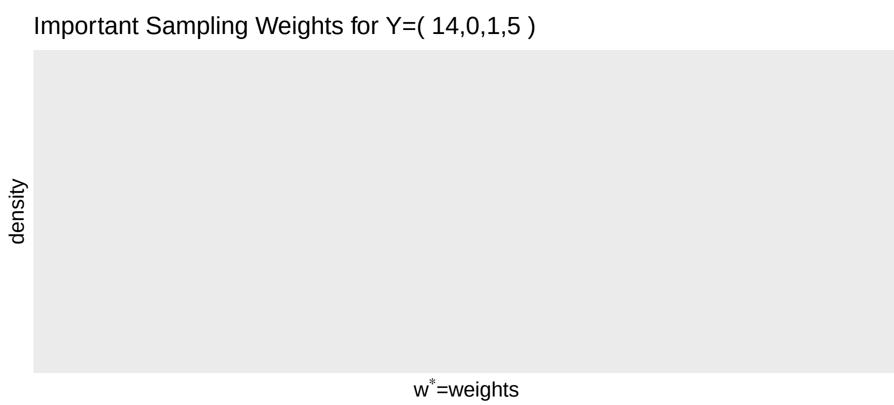
$$\text{Normal Approximation for } Y = (125, 18, 20, 34) \sim \mathcal{N}(\mu = 0.90344,\ \sigma = 0.09348)$$



	IS	Norm.Apprx	Diff
mean	NaN	0.90344	NaN
sd	NaN	0.09348	NaN



	IS	Norm.Apprx	Diff
mean	NaN	0.90344	NaN
sd	NaN	0.09348	NaN



	IS	Norm.Apprx	Diff
mean	NaN	0.90344	NaN
sd	NaN	0.09348	NaN

Problem 3c

Repeat (a) and (b) with a Uniform[0, 1] importance function.

Problem 4

Problem 4a

Solve the following problem posted by the Reverend Thomas Bayes in his essay "Essay Towards Solving a Problem in the Doctrine of Chances," which was published in the *Philosophical Transactions of the Royal Society* (London) in 1763: *Given* the number of times in which an unknown event has happened and failed: *Required* the chance that the probability of its happening in a single trial lies somewhere between any tow degrees of probability that can be named.

In other words, if the number of the successful happenings of the event is p and the failures q, and if the named "degrees" of the probability are b and f, respectively, compute:

$$\int_b^f x^p(1-x)^q dx \Bigg/ \int_0^1 x^p(1-x)^q dx$$

via important sampling. Take $p = 1,\ q = 5,\ b = 0.7,\ f = 0.9$.

Problem 4b

Repeat the calculation using numerical integration. Compare the results of (a) and (b).

Problem 6a

Under the likelihood $\theta^k(1-\theta)^{n-x}$ and the Beta(a, b) prior (a and b known) compute the exact posterior mean. Repeat the calculation using the second-order Laplace approximation. evaluate the relative error for the data $n = 5,\ x = 3$ and the prior values $a = b = 1/2$. What is the relative error when $n = 25,\ x = 15$ (same prior)?

Problem 1

Recall the genetic linkage model of Section 4.1.

Problem 1a

For the data $Y = (125, 18, 202, 34)$ implement the *EM* algorithm. Use a flat prior on θ. Try starting your algorithm at $\theta = .1, .2, .3, .4, .6$ and .8. Did the algorithm converge for all of these starting values? How do you access convergence? How many iterations were required for convergence?

Problem 1c

Plot the normal approximation along with the normalized likelihood. Is the normal approximation appropriate in this case?

Problem 1d

Repeat (a) and (c) for the data $Y = (14, 0, 1, 5)$. did the algorithm coverage for all of the above starting values?

Problem 2

Repeat Problem 1 (a) and (d) using the Monte Carlo *EM*. How did you assess convergence.