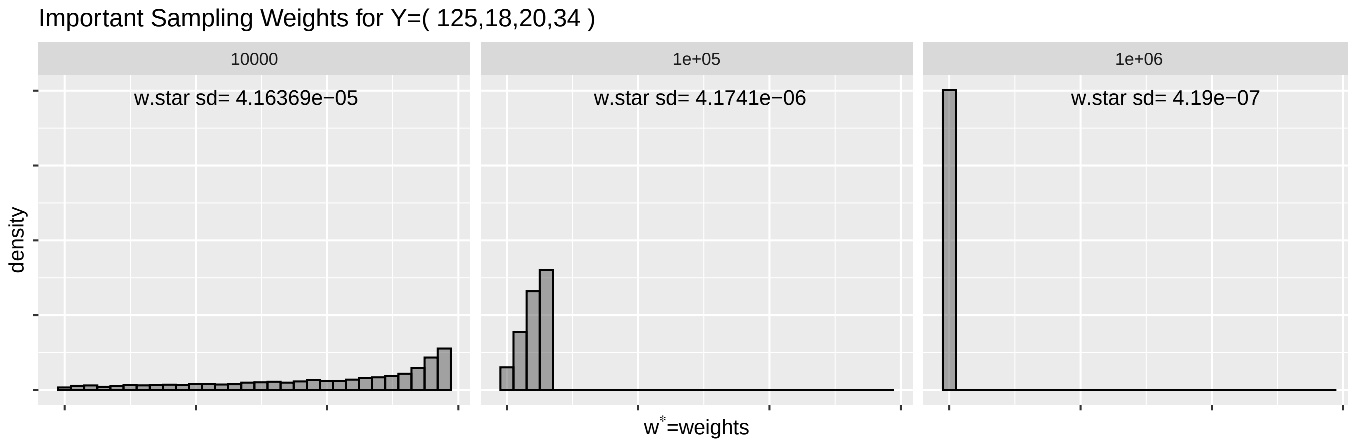
STAT 457 Homework 05

Martha Eichlersmith

2019-12-03

Problem 3a

For the genetic linkage model: 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. Normal Approxiamation for $Y = (125, 18, 20, 34) \sim \mathcal{N}(\mu = 0.62682,\ \sigma = 0.05382)$



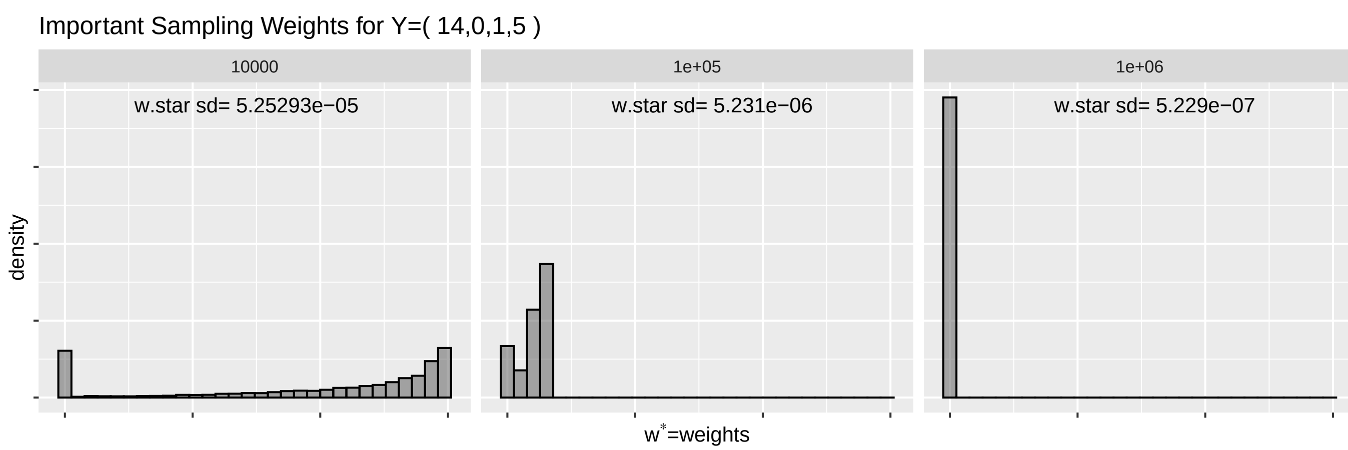
	IS	Norm.Apprx	Diff
mean	0.62365	0.62682	−0.00317
sd	0.03715	0.05382	−0.01667

	IS	Norm.Apprx	Diff
mean	0.62338	0.62682	−0.00344
sd	0.03706	0.05382	−0.01676

	IS	Norm.Apprx	Diff
mean	0.62347	0.62682	−0.00335
sd	0.03716	0.05382	−0.01666

Problem 3b

Repeat (a) for the data $Y = (14, 0, 1, 5)$. Normal Approximation for $Y = (125, 18, 20, 34) \sim \mathcal{N}(\mu = 0.90344,\ \sigma = 0.09348)$



	IS	Norm.Apprx	Diff
mean	0.87758	0.90344	−0.02586
sd	0.06194	0.09348	−0.03154

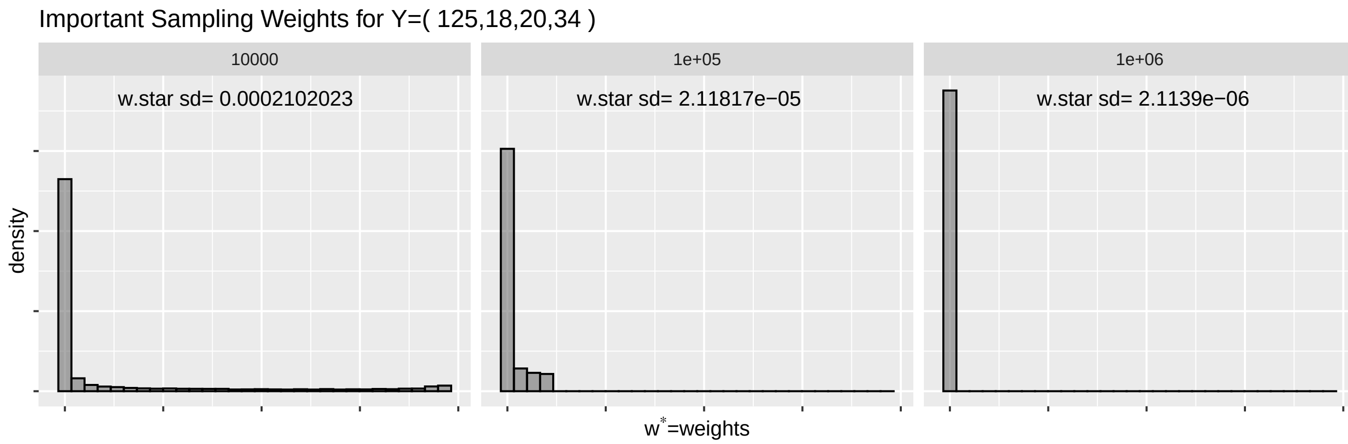
	IS	Norm.Apprx	Diff
mean	0.87714	0.90344	−0.0263
sd	0.06214	0.09348	−0.03134

	IS	Norm.Apprx	Diff
mean	0.87713	0.90344	−0.02631
sd	0.0622	0.09348	−0.03128

Using a normal important sampling function to estiamte the posterior mean is closer for $Y = (125, 18, 20, 34)$ normal approximation than $Y = (14, 0, 1, 5)$. This makes sense as in the last homework, we showed the likelihood for the first data follows the approximate normal distribution very closely whereas the second data likelihood did not follow the normal approximation well.

Problem 3c

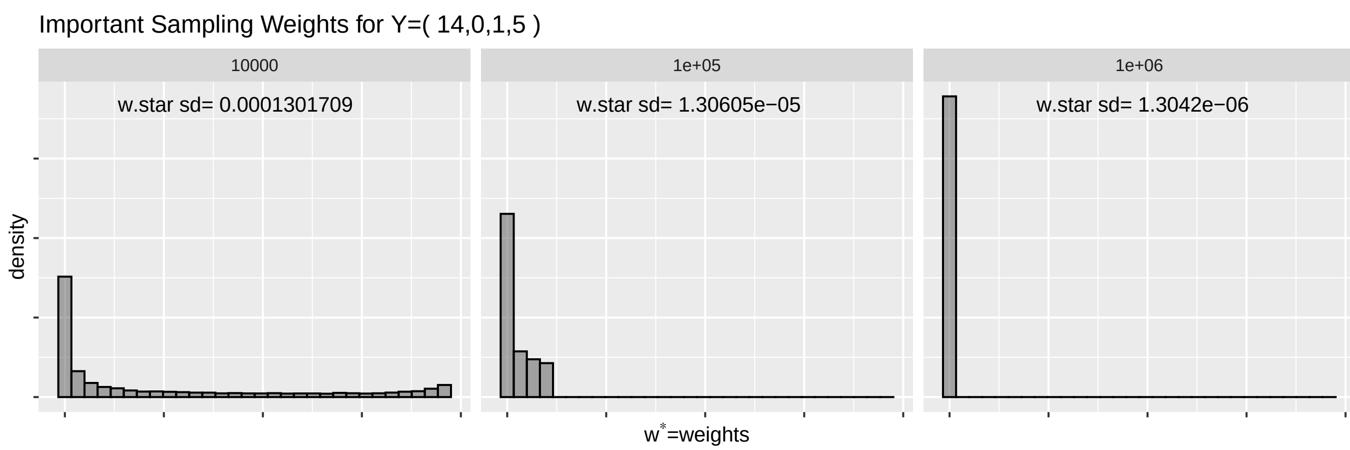
Repeat (a) and (b) with a Uniform[0, 1] importance function.



	IS	Norm.Apprx	Diff
mean	0.61867	0.90344	−0.28477
sd	0.05209	0.09348	−0.04139

	IS	Norm.Apprx	Diff
mean	0.61838	0.90344	−0.28506
sd	0.05169	0.09348	−0.04179

	IS	Norm.Apprx	Diff
mean	0.61855	0.90344	−0.28489
sd	0.05143	0.09348	−0.04205



	IS	Norm.Apprx	Diff
mean	0.81241	0.90344	−0.09103
sd	0.11825	0.09348	0.02477

	IS	Norm.Apprx	Diff
mean	0.81389	0.90344	−0.08955
sd	0.11782	0.09348	0.02434

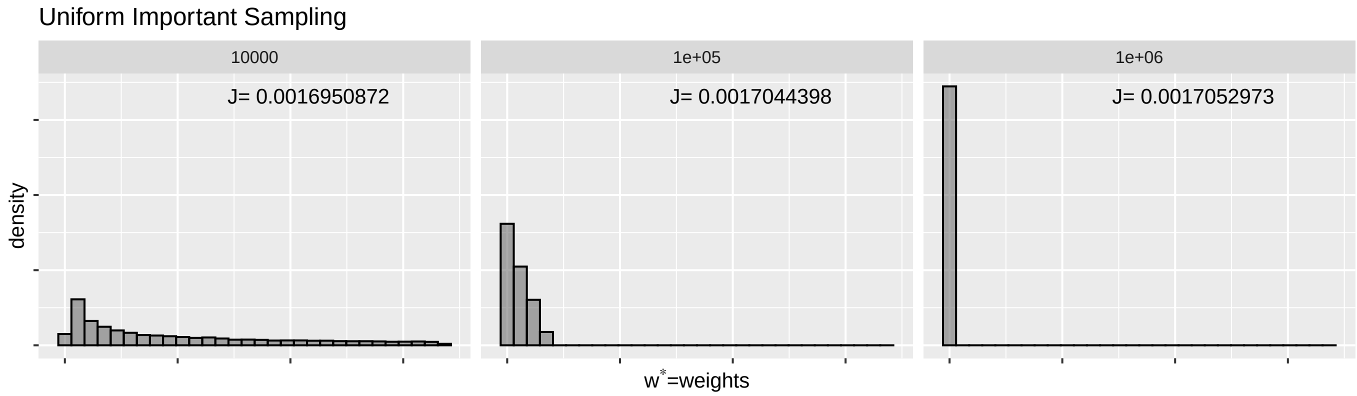
	IS	Norm.Apprx	Diff
mean	0.81412	0.90344	−0.08932
sd	0.11791	0.09348	0.02443

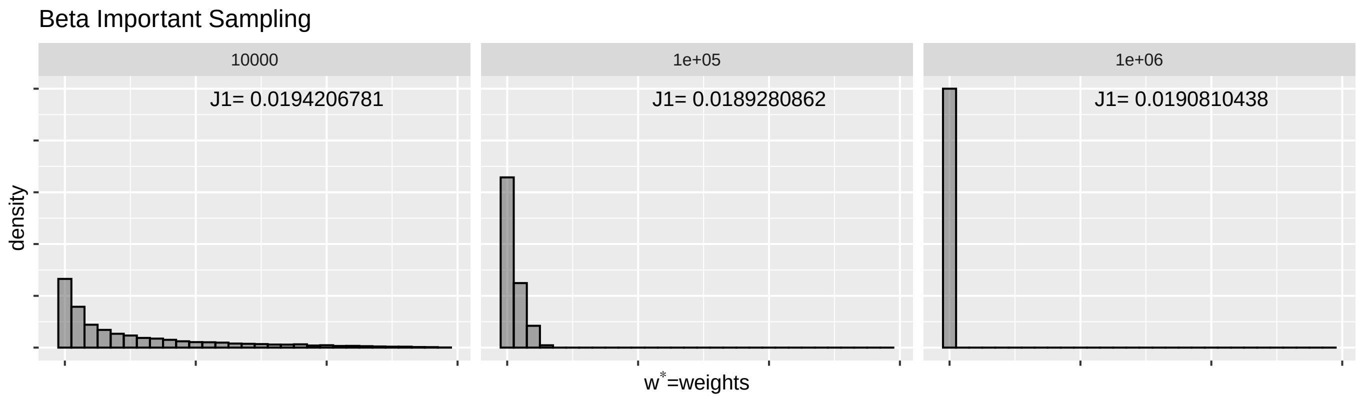
Note that the histograms of the weights, w^*, are very similar for both sets of data. This is because the importance function is not dependent on the data (like it was for when using the normal approximation data for a norma importance function).

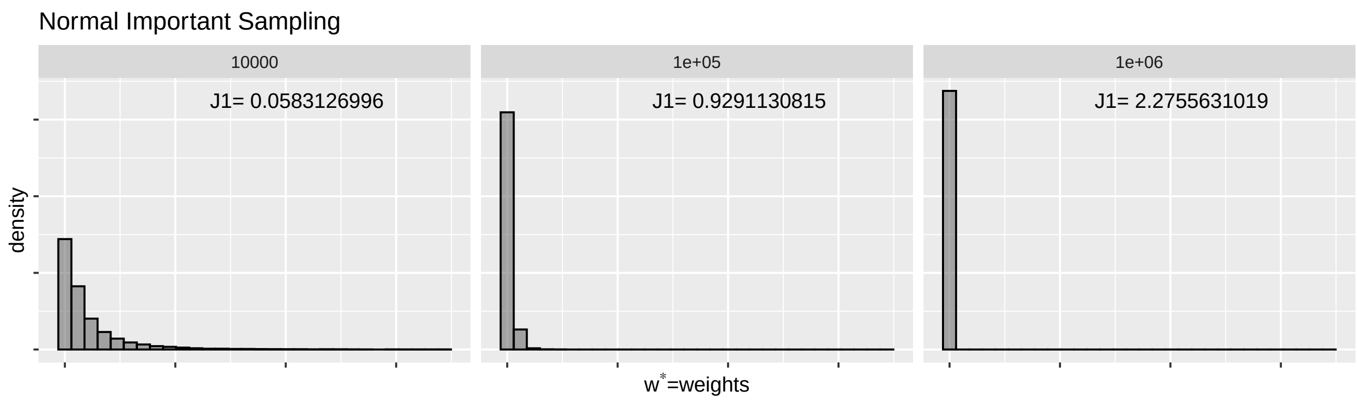
Problem 4

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: $\int_b^f x^p(1-x)^q dx / \int_0^1 x^p(1-x)^q dx$ via important sampling. Take $p = 1,\ q = 4,\ b = 0.7,\ f = 0.9$.







Problem 4b

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

	Integration	IS - Uniform			IS - Beta			IS- Normal		
It	N/A	10000	1e+05	1e+06	10000	1e+05	1e+06	10000	1e+05	1e+06
J	0.01088	0.0017	0.0017	0.00171	0.01942	0.01893	0.01908	0.05831	0.92911	2.27556
J1-Intg	0	-0.00918	-0.00918	-0.00917	0.00854	0.00805	0.0082	0.04743	0.91823	2.26468

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)?

Exact

$$p_E(\theta \mid Y) \propto \text{Beta}(x+a,\ n-x+b) \quad \Longrightarrow \quad \mu_E = \frac{x+a}{a+b+n}$$

2nd Laplace Approx

$$\mu_L = \frac{\sigma^*}{\sigma^\dagger} \cdot \frac{\exp\{-nh^*(\theta^*)\}}{\exp\{-nh^\dagger(\theta^\dagger)\}}$$

$$-nh^\dagger(\theta) = \ell(\theta \mid Y) + \ln(p(\theta))$$

$$\ell(\theta \mid Y) = x\ln(\theta) + (n-x)\ln(\theta)$$

$$\ln(p(\theta)) = (a-1)\ln(\theta) + (b-1)\ln(1-\theta)$$

$$= \alpha_\dagger \ln(\theta) + \beta_\dagger \ln(1-\theta) \qquad \begin{aligned} \alpha_\dagger &= x+a-1 \\ \beta_\dagger &= n-x+b-1 \end{aligned}$$

$$-nh^*(\theta) = \ell(\theta \mid Y) + \ln(p(\theta)) + ln(g(\theta))$$

$$\begin{aligned} \ell(\theta \mid Y) &= x\ln(\theta) + (n-x)\ln(\theta) \\ \ln(p(\theta)) &= (a-1)\ln(\theta) + (b-1)\ln(1-\theta) \\ ln(g(\theta)) &= \ln(\theta) \end{aligned}$$

$$= \alpha_* \ln(\theta) + \beta_* \ln(1-\theta) \qquad \begin{aligned} \alpha_* &= x+a \\ \beta_* &= n-x+b-1 \end{aligned}$$

$$\theta^{(\cdot)} = \arg\max_\theta(-nh^{(\cdot)}(\theta)) = \frac{\partial - nh^{(\cdot)}(\theta)}{\partial\theta} = \frac{\alpha_{(\cdot)}}{\theta} - \frac{\beta_{(\cdot)}}{1-\theta} \overset{\text{set}}{=} 0 \implies \theta^{(\cdot)} = \frac{\alpha_{(\cdot)}}{\alpha_{(\cdot)} + \beta_{(\cdot)}}$$

$$\sigma^{(\cdot)} = \left[\frac{\partial^2 h^{(\cdot)}(\theta)}{\partial\theta^2}\bigg|_{\theta^{(\cdot)}}\right]^{-1/2} = \left[\frac{1}{n}\left(\frac{\alpha_{(\cdot)}}{(\theta^{(\cdot)})^2} + \frac{\beta_{(\cdot)}}{(1-\theta^{(\cdot)})^2}\right)\right]^{-1/2}$$

	Part.1	Part.2
Data.n	5	25
Data.x	3	15
Prior.a	0.5	0.5
Prior.b	0.5	0.5
Exact.mean	0.58333	0.59615
Laplace.Mean	0.78538	0.62949
Relative.Error	0.34636	0.05591

Problem 1

Recall the genetic linkage model of Section 4.1.

Problem 1a

For the data $Y = (125, 18, 20, 34)$ 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?

$$\text{Given } Y = (y_1, y_2, y_3, y_4) \text{ with probabilities } \left(\frac{\theta+2}{4}, \frac{1-\theta}{4}, \frac{1-\theta}{4}, \frac{\theta}{4}\right)$$
$$\text{Say } Y = (x, z, y_2, y_3, y_4) \text{ with probabilities } \left(\frac{1}{2}, \frac{\theta}{4}, \frac{1-\theta}{4}, \frac{1-\theta}{4}, \frac{\theta}{4}\right)$$

$$p(Z \mid Y, \theta) \sim \text{Binomial}\left(x+z,\ \frac{p(z)}{p(x)+p(z)}\right) = \text{Binomial}\left(y_1,\ \frac{\theta/4}{(\theta+2)/4}\right) = \text{Binomial}\left(y_1,\ \frac{\theta}{2+\theta}\right)$$

$$p(\theta \mid Y, Z) \sim \left(\frac{1}{2}\right)^x \left(\frac{\theta}{4}\right)^z \left(\frac{1-\theta}{4}\right)^{y_2+y_3} \left(\frac{\theta}{4}\right)^{y_4} \propto \theta^{z+x_5}(1-\theta)^{y_2+y_3}$$

$$\underline{\textbf{E Step}: \text{to get Q function}}$$
$$\begin{aligned} Q(\theta, \theta^i) &= \mathbb{E}_{Z|\theta^i}\left[\log\left(p(\theta \mid Y, Z)\right)\right] = \mathbb{E}_{Z|\theta^i}\Big[(z+y_4)\log(\theta) + (y_2+y_3)\log(1-\theta) \mid \theta^i, Y\Big] \\ &= (y_2+y_3)\log(1-\theta) + \left(\mathbb{E}_{Z|\theta^i}\left[Z \mid \theta^i, Y\right] + y_4\right)\log(\theta) \end{aligned}$$
$$p(Z \mid Y, \theta^i) \sim \text{Binomial}(y_1, \theta/(\theta+2)) \implies \mathbb{E}_{Z|\theta^i}\left[Z \mid \theta^i, Y\right] = \frac{y_1\theta^i}{\theta^i+2}$$
$$= (y_2+y_3)\log(1-\theta) + \left(\frac{y_1\theta^i}{\theta^i+2} + y_4\right)\log(\theta)$$

$$\underline{\textbf{M Step}: \arg\max_\theta Q(\theta, \theta^i)}$$
$$\frac{\partial Q(\theta, \theta^i)}{\partial \theta} = -\frac{y_2+y_3}{1-\theta} + \frac{\mathbb{E}Z}{\theta} \overset{\text{set}}{=} 0 \qquad \implies \theta^{i+1} = \frac{\mathbb{E}Z + y_4}{\mathbb{E}Z + y_2 + y_3 + y_4}$$
$$\theta^{i+1} = \frac{\frac{y_1\theta^i}{\theta^i+2} + y_4}{\frac{y_1\theta^i}{\theta^i+2} + y_2 + y_3 + y_4}$$

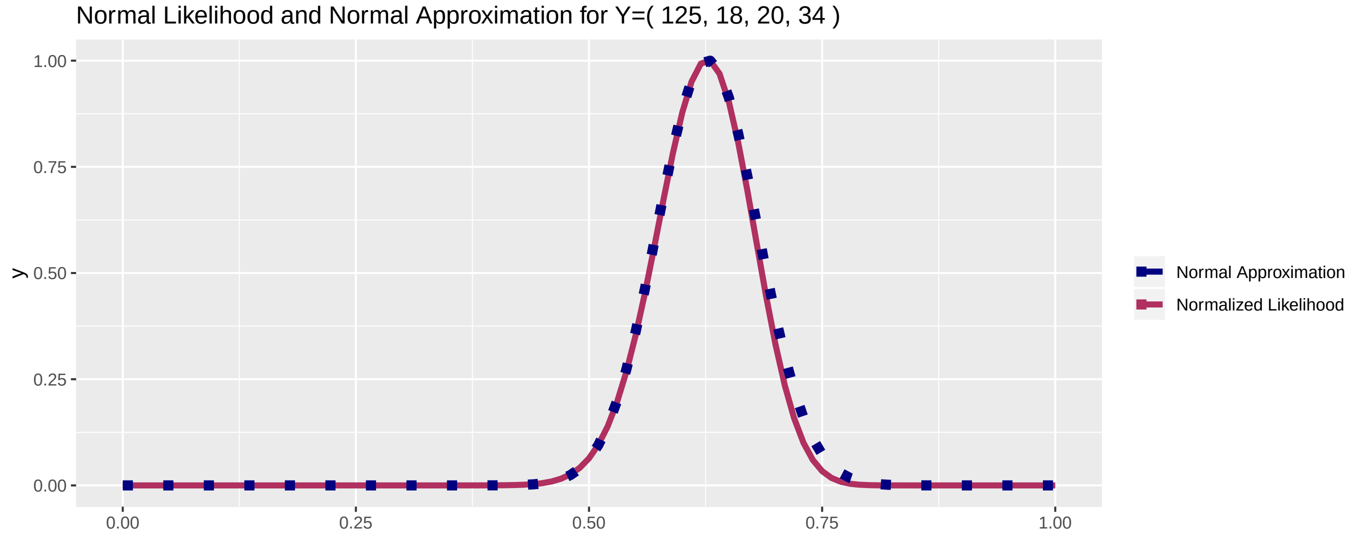
EM for Y=(125,18,20,34)

Start Value	0.1	0.2	0.3	0.4	0.6	0.8
Estimation	0.6268214856	0.6268214893	0.6268214921	0.6268214943	0.6268214952	0.6268215125
Iterations Used	10	10	10	10	9	9

Convergence is determined if the the values within the chain have an absolute difference less than 1e-07.

Problem 1c

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



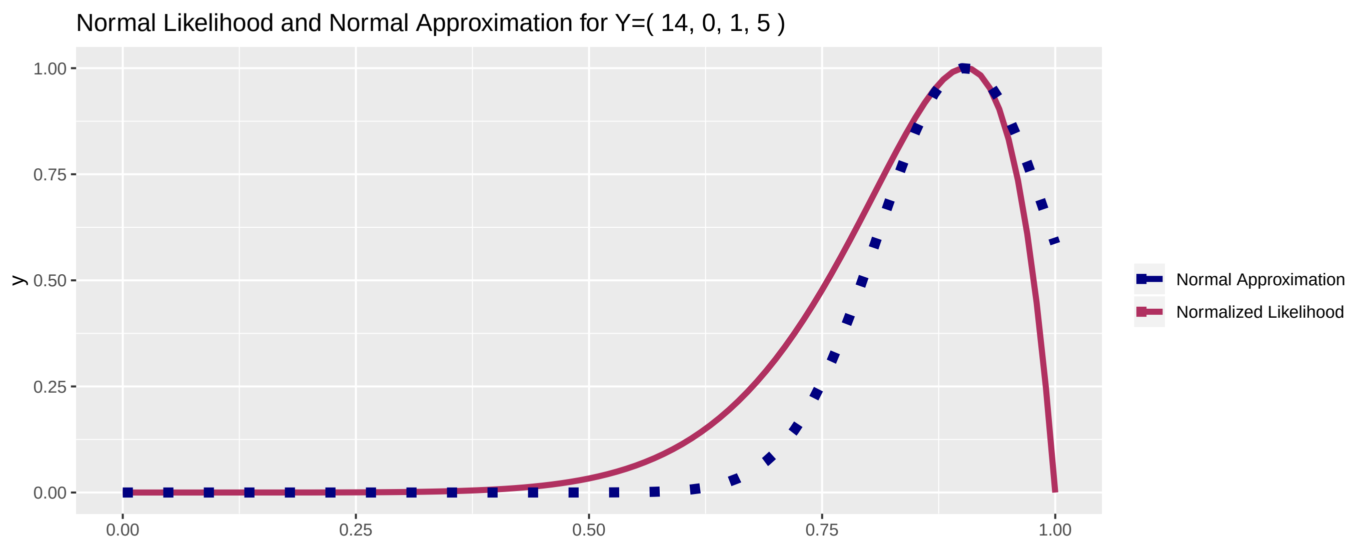
The Normal Approximation is appropirate in this case.

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?

EM for Y=(14,0,1,5)

Start Value	0.1	0.2	0.3	0.4	0.6	0.8
Estimation	0.9034401126	0.9034401130	0.9034401133	0.9034401135	0.9034401139	0.9034401110
Iterations Used	7	7	7	7	7	6



The Normal Approximation is not appropriate in this case.

Problem 2

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

MCEM for Y=(125,18,20,34), m= 100

Start Value	0.1	0.2	0.3	0.4	0.6	0.8
Estimation	0.6264622039	0.6277429467	0.6288700068	0.6267923787	0.6253573893	0.6275240149
Iterations Used	340	186	50	164	76	144

MCEM for Y=(125,18,20,34), m= 1000

Start Value	0.1	0.2	0.3	0.4	0.6	0.8
Estimation	0.6267887133	0.6271292880	0.6275057590	0.6265613176	0.6268693355	0.6265429672
Iterations Used	464	589	375	2415	541	464

MCEM for Y=(125,18,20,34), m= 10000

Start Value	0.1	0.2	0.3	0.4	0.6	0.8
Estimation	0.6268579772	0.6267718515	0.6267678191	0.6268451525	0.6269261166	0.6267568213
Iterations Used	148	937	2737	331	763	1362

MCEM for Y=(14,0,1,5), m= 100

Start Value	0.1	0.2	0.3	0.4	0.6	0.8
Estimation	0.9044890162	0.9032882012	0.9043977055	0.9028182702	0.9040307102	0.9038461538
Iterations Used	101	28	34	155	171	93

MCEM for Y=(14,0,1,5), m= 1000

Start Value	0.1	0.2	0.3	0.4	0.6	0.8
Estimation	0.9029031945	0.9031195505	0.9034562657	0.9035772828	0.9031945789	0.9034189685
Iterations Used	199	113	521	151	73	94

MCEM for Y=(14,0,1,5), m= 10000

Start Value	0.1	0.2	0.3	0.4	0.6	0.8
Estimation	0.9033881766	0.9034413503	0.9034441473	0.9034553336	0.9033059689	0.9032376676
Iterations Used	330	551	152	359	1304	367