STA 360: Homework 2

Samuel Eure

9/12/2018

Problem 3.1

(3.1a)

Given the set up of the problem it seem to be appropriate to model Y as a random variable sampled from a Bernuli distribution with parameter θ. Thus,

$$P(Y_1 = y_1, ..., Y_{100} = dy_{100} \mid \theta) = \Pi_{i=1}^{100} P(Y_i = y_i \mid \theta) = \Pi_{i=1}^{100} \theta^{y_i}(1-\theta)^{1-y_i} = \theta^{\sum_{i=1}^{n} y_i}(1-\theta)^{n-\sum_{i=1}^{n} y_i}$$

$$P(Y_1 = y_1, ..., Y_{100} = dy_{100} \mid \theta) = \theta^{\sum_{i=1}^{n} y_i}(1-\theta)^{n-\sum_{i=1}^{n} y_i}$$

This differs from the $P(\sum_{i=1}^{n} Y_i \mid \theta)$, which takes the form of a Binomial Distribution since the sum of i.i.d. Bernuli random variables is binomial. Thus,

$$P(\sum_{i=1}^{n} Y_i = y_i \mid \theta) = \binom{n}{\sum_{i=1}^{n} y_i} \theta^{\sum_{i=1}^{n} y_i}(1-\theta)^{n-\sum_{i=1}^{n} y_i}$$

(3.1b)

Using the formulas

```
thetas = seq(0.0, 1.0, 0.1)
Y = 57 #Given from the problem
Py_t = dbinom(57, 100, thetas)
for (t in thetas){
  print(paste("Pr(sum(Y)=57|theta=",t,")=",Py_t[t*10+1]))
}
```

```
## [1] "Pr(sum(Y)=57|theta= 0 )= 0"
## [1] "Pr(sum(Y)=57|theta= 0.1 )= 4.10715691923436e-31"
## [1] "Pr(sum(Y)=57|theta= 0.2 )= 3.73845866856749e-16"
## [1] "Pr(sum(Y)=57|theta= 0.3 )= 1.30689479922164e-08"
## [1] "Pr(sum(Y)=57|theta= 0.4 )= 0.000228579175709522"
## [1] "Pr(sum(Y)=57|theta= 0.5 )= 0.0300686426442146"
## [1] "Pr(sum(Y)=57|theta= 0.6 )= 0.0667289496743285"
## [1] "Pr(sum(Y)=57|theta= 0.7 )= 0.00185317154808753"
## [1] "Pr(sum(Y)=57|theta= 0.8 )= 1.00353485743407e-07"
## [1] "Pr(sum(Y)=57|theta= 0.9 )= 9.39585764212832e-18"
## [1] "Pr(sum(Y)=57|theta= 1 )= 0"
```

```
plot(thetas, Py_t, type = 'o', col = "blue", lwd = 2,
     xlab = "Theta", ylab="P(sum[Y]=57| Theta)",
     main = "Problem 3.1b Plot")
```

Problem 3.1b Plot
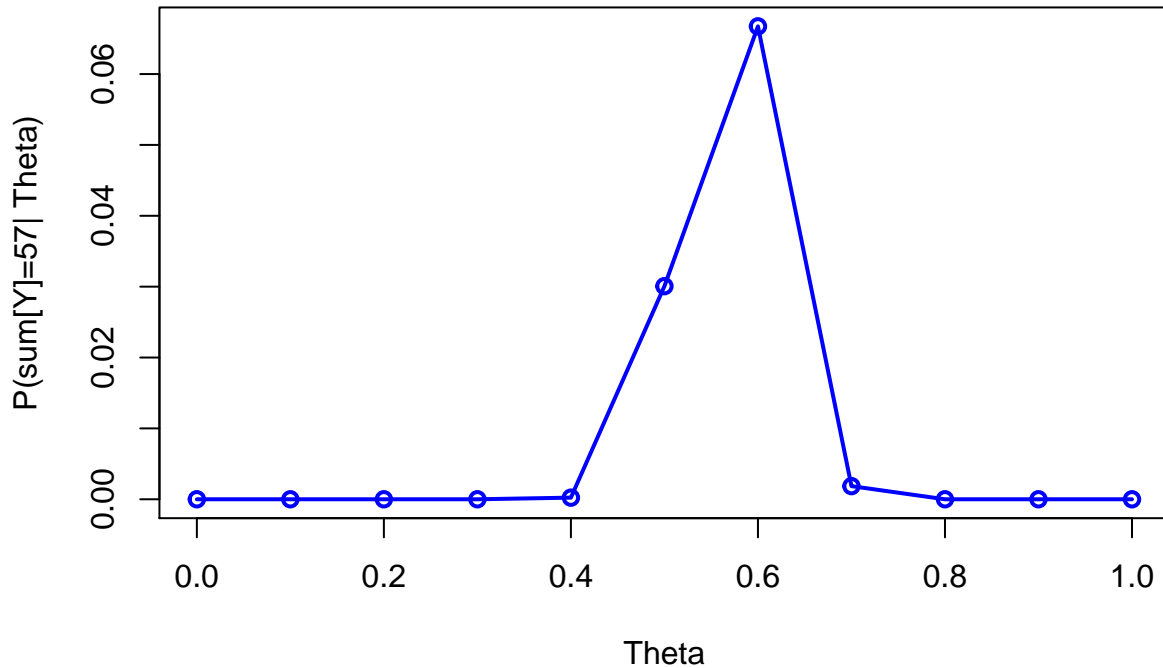
(3.1c)

Suppose $P(\theta) = \frac{1}{11}$ is our prior belief of θ, where $\theta \in 0.0, 0.1, ..., 1.0$. Using Bayes' rule,

$$p(\theta \mid \sum_{i=1}^{n} Y_i = 57) = \frac{P(\sum_{i=1}^{n} Y_i = 57 \mid \theta)P(\theta)}{P(\sum_{i=1}^{n} Y_i = 57)} = \frac{P(\sum_{i=1}^{n} Y_i = 57 \mid \theta)P(\theta)}{\sum_{k=1}^{11} P(\sum_{i=1}^{n} Y_i = 57 \mid \theta_k)P(\theta_k)} = \frac{P(\sum_{i=1}^{n} Y_i = 57 \mid \theta)\frac{1}{11}}{\sum_{k=1}^{11} P(\sum_{i=1}^{n} Y_i = 57 \mid \theta_k)\frac{1}{11}}$$

$$= \frac{\frac{1}{11}P(\sum_{i=1}^{n} Y_i = 57 \mid \theta)}{\frac{1}{11}\sum_{k=1}^{11} P(\sum_{i=1}^{n} Y_i = 57 \mid \theta_k)} = \frac{P(\sum_{i=1}^{n} Y_i = 57 \mid \theta)}{\sum_{k=1}^{11} P(\sum_{i=1}^{n} Y_i = 57 \mid \theta_k)}$$

I will compute $p(\theta \mid \sum_{i=1}^{100} Y_i = 57)$ by simply dividing the values of $P(\sum_{i=1}^{n} Y_i = 57 \mid \theta)$ from 3.1b by the sum of $P(\sum_{i=1}^{n} Y_i = 57 \mid \theta)$

```
D = rep(sum(Py_t),11)
Pt_y = Py_t/D
for (t in thetas){
  print(paste("Pr(theta=",t,"| sum(Y)=57",")=",Py_t[t*10+1]))
}
```

```
## [1] "Pr(theta= 0 | sum(Y)=57 )= 0"
## [1] "Pr(theta= 0.1 | sum(Y)=57 )= 4.10715691923436e-31"
## [1] "Pr(theta= 0.2 | sum(Y)=57 )= 3.73845866856749e-16"
## [1] "Pr(theta= 0.3 | sum(Y)=57 )= 1.30689479922164e-08"
## [1] "Pr(theta= 0.4 | sum(Y)=57 )= 0.000228579175709522"
## [1] "Pr(theta= 0.5 | sum(Y)=57 )= 0.0300686426442146"
## [1] "Pr(theta= 0.6 | sum(Y)=57 )= 0.0667289496743285"
## [1] "Pr(theta= 0.7 | sum(Y)=57 )= 0.00185317154808753"
```

```
## [1] "Pr(theta= 0.8 | sum(Y)=57 )= 1.00353485743407e-07"
## [1] "Pr(theta= 0.9 | sum(Y)=57 )= 9.39585764212832e-18"
## [1] "Pr(theta= 1 | sum(Y)=57 )= 0"
```

```
plot(thetas, Pt_y, type = 'o', lwd = 2, col = 'red',
     xlab = "Theta", ylab = "P(Theta | sum[Y]=57)", main = "Problem 3.1c Plot")
```
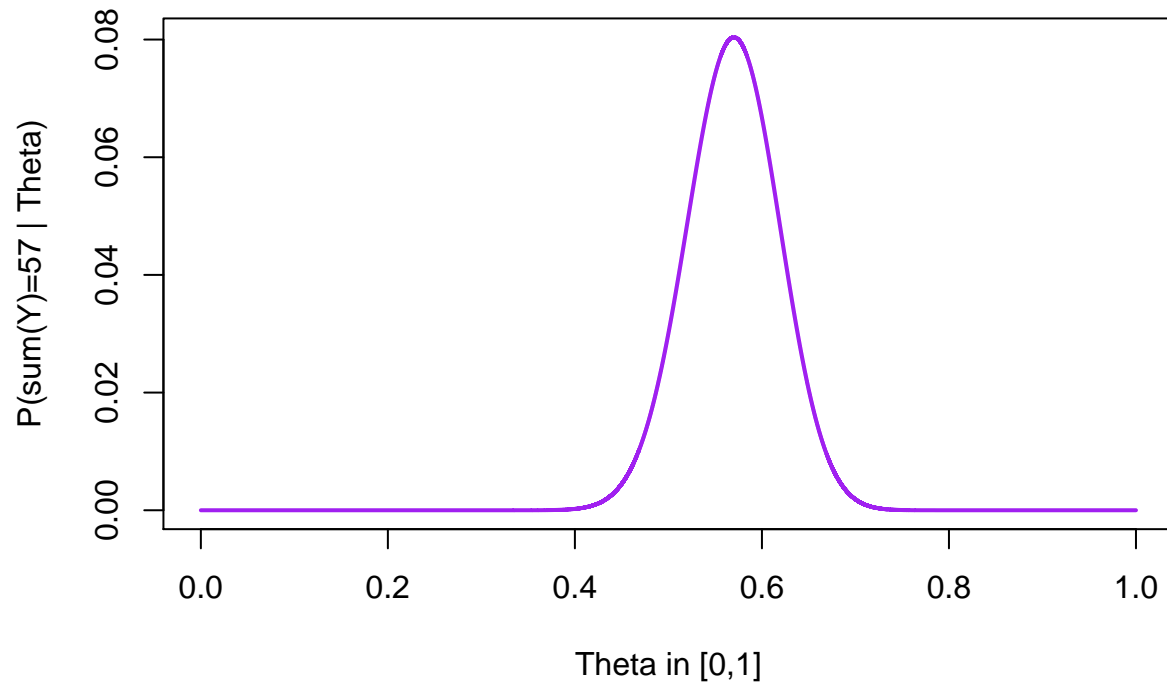
Problem 3.1c Plot
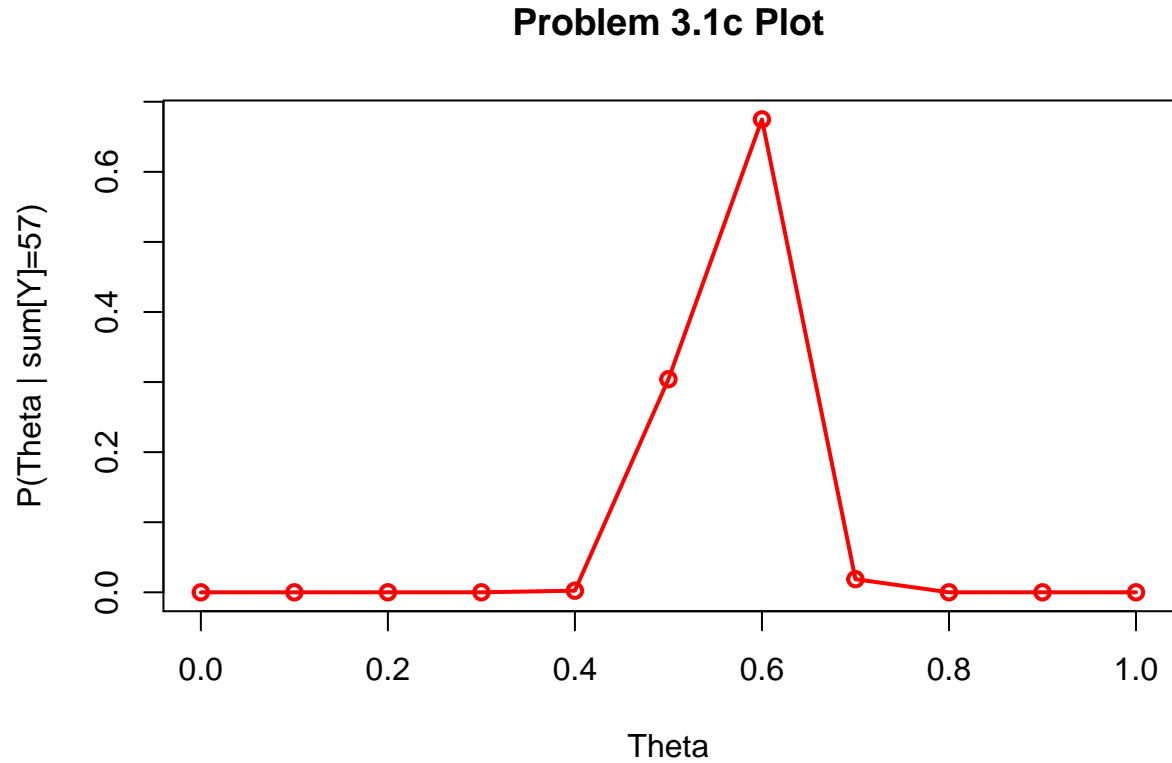
3.1d

Now suppose $\theta \in [0,1] \subset \mathbb{R}$ and $p(\theta) = 1, \forall \theta$.

```
continuousThetas = seq(0.0, 1.0, 0.00001)
Py_continuousTheta = dbinom(57, 100, continuousThetas)*1 #Multiplied by 1 for p(theta) = 1
plot(continuousThetas, Py_continuousTheta, type = 'l',
     col = 'purple', lwd = 2, xlab = "Theta in [0,1]", ylab="P(sum(Y)=57 | Theta)",
     main = "Problem 3.1d Plot")
```
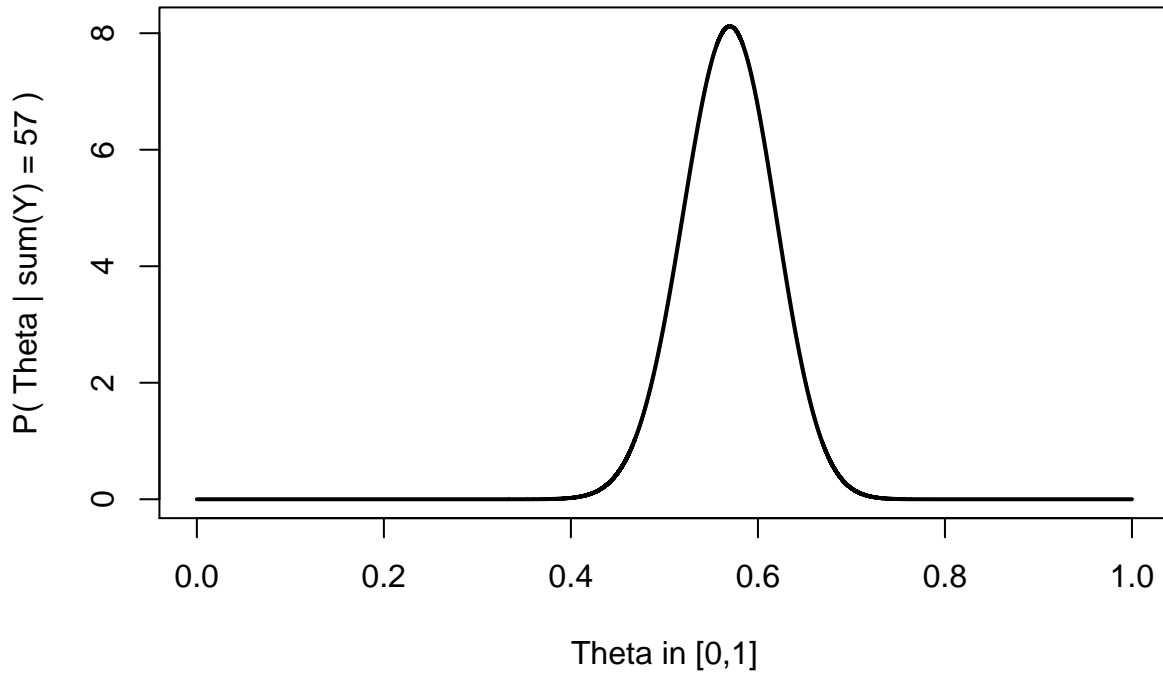
Problem 3.1d Plot

3.1e

Using a beta(1,1) as my prior distribution of θ and a posterior of beta(1+57, 1 + 100-57) = beta(58, 44),

$$P(\sum_{i=1}^{100} Y_i = 57) = P(\sum_{i=1}^{100} Y_i = 57 \mid \theta)P(\theta \mid)$$

```
PostTheta = dbeta(continuousThetas, 1+57, 1+100-57)
plot(continuousThetas, PostTheta, type = 'l', lwd = 2, xlab="Theta in [0,1]",
     ylab="P( Theta | sum(Y) = 57 )", main = "Problem 3.1e Plot")
```

Problem 3.1e Plot

Comments about the plots of problem 3.1

As the grader may have noticed by now, all the plots have the same general shape. To be more specific, the plots from **3.1b** and **3.1c** have identical shapes and differ by a constant scaling factor, while plots **3.1d** and **3.1e** have identical shapes and differ by a constant scaling factor. This makes sense since the uniform distribution and beta distribution (which in the continuous case are the same thing) are a conjugate priors to the binomial distribution. On additional comparison between the plots is that the plots in **3.1b&c** aren't as smooth as the plots in **3.1d&e**, and it appears that our choice of a distrete prior in problem **3.1b&c** may not have been a good choice since it results in an awkward posterior distribution.

Problem 3.2

Given $\theta_0 = \frac{a}{a+b}$, $n_0 = a + b$, **with** $\theta \in 0.1, 0.2, ..., 0.9$ **and** $n_0 \in 1, 2, 8, 16, 32$, **compute** $P(\theta > 0.5 \mid \sum Y_i = 57)$ **with a beta(a,b) prior on** θ.

With this set up, $(\theta \mid \sum Y_i = 57) \sim$ **beta(a+57, b+100-57)**

```
Thetas = seq(0.1, 0.9, 0.1)
N      = c(1, 2, 8, 16, 32)
P = c()
for (t in Thetas){
  rowT = c()
  for (n in N){
    a = t*n
    b = (1-t)*n
```

```
    p = 1- pbeta(0.5, a + 57, b + 100-57)
    rowT = c(rowT,p)
  }
  P = c(P, rowT)
}
P = matrix(P, nrow = length(Thetas), ncol = length(N), byrow = TRUE)
cols = rev(colorRampPalette(c('black','red',"yellow",'white'))(20))

filled.contour(Thetas, N, P, col =cols, main = 'Contour Plot of P(Theta > 0.5 | Sum(Y) = 57)',
               xlab = "Theta", ylab = "n", plot.axes = {axis(1, Thetas)
                  axis(2, seq(0, 32, 2))}, nlevels = 20)
```

Contour Plot of P(Theta > 0.5 | Sum(Y) = 57)
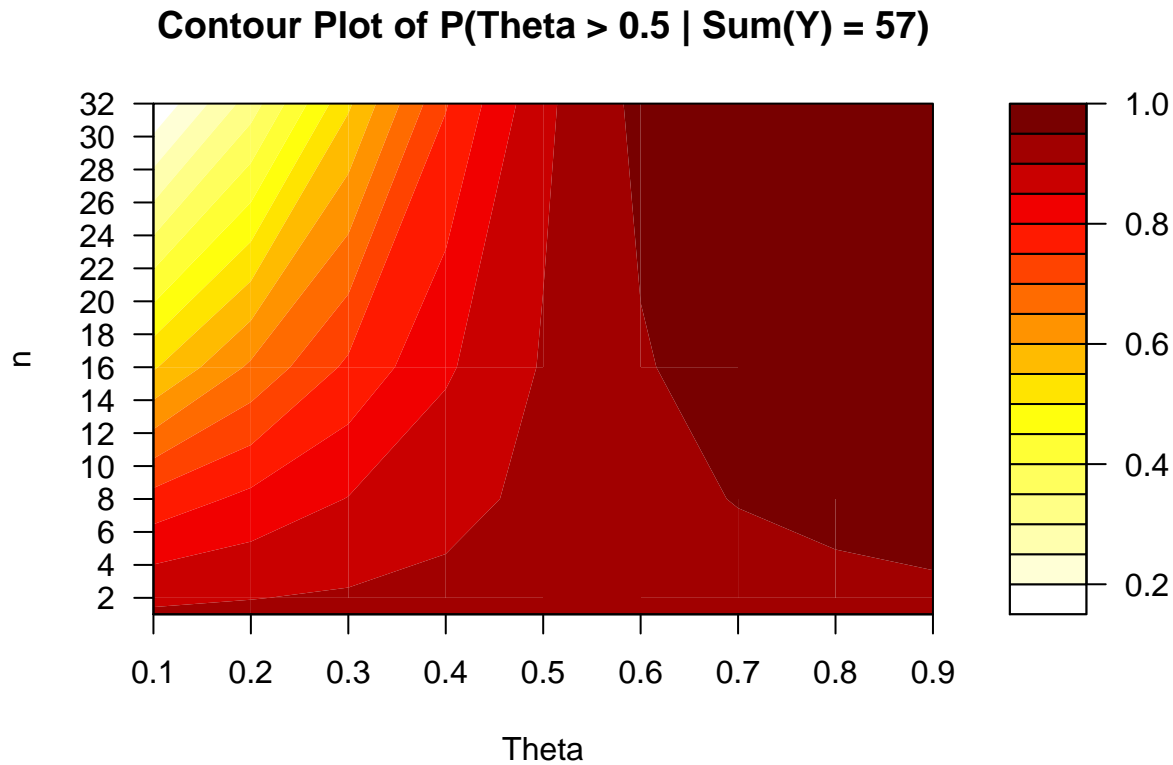
The plot above shows the probability that $\theta > 0.5$ by color given that $\sum Y_i = 57$. The x-axis shows what our prior belief of θ was while the y-axis shows the "weight" of our prior belief – how strongly we felt about our prior belief. This plot shows that people who didn't feel very strongly at all about any particular θ (which correspondes to $n < 10 orso$), there is a decent chance (probability > 0.7ish) that $\theta > 0.5$ given the data; this is evidenced by the red coloring of the plot below and at the $n = 10$ horizen. Moreover, is someone originally believed that $\theta \in [0.4, 0.9]$, then they should be very confident that $\theta > 0.5$ given the data. The only people who should remain skeptical that $\theta > 0.5$ are those who were originally very confident $(n \in [10, 31])$ that $\theta \in [0.1, 0.25]$.

Proof of identity on page 33

Prove:

$$\text{beta}(a, b) = \int_0^1 x^{a-1}(1-x)^{b-1}dx = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

6

Proof: First, let me do some clever substititions. Let $n = a + b$. Then,

$$\int_0^1 x^{a-1}(1-x)^{b-1}dx = \int_0^1 x^{a-1}(1-x)^{n-a-1}dx$$

Now, I need to tackle this integral by parts. Let $f(x) = x^{a-1}$ and $g(x) = (1-x)^{b-1}$. Then,

$$\int_0^1 x^{a-1}(1-x)^{n-a-1}dx = \int_0^1 f(x)g(x)dx = [f(x)G(x)]_0^1 - \int_0^1 f'(x)G(x)dx$$

Where $G(x)$ is the integrade of $g(x)$ and $f'(x) = \frac{d}{dx}f(x)$. Then,

$$\int_0^1 x^{a-1}(1-x)^{b-1}dx = \left[x^{a-1}(n-a-1)(1-x)^{(n-a-1)-1} \right]_0^1 - \int_0^1 \frac{x^{(a-1)+1}}{(a-1)+1}(n-a-1)(-1)(1-x)^{(n-a-1)-1}dx$$

Since

$$\left[x^{a-1}(n-a-1)(1-x)^{(n-a-1)-1} \right]_0^1 = [1^{a-1}(n-a-1)(1-1)^{(n-a-1)-1}] - [0^{a-1}(n-a-1)(1-0)^{(n-a-1)-1}] = 0 - 0 = 0$$

$$\implies$$

$$\int_0^1 x^{a-1}(1-x)^{b-1}dx = -\int_0^1 \frac{x^a}{(a-1)+1}(n-a-1)(-1)(1-x)^{(n-a-1)-1}dx = \frac{(n-a-1)}{(a-1)+1}\int_0^1 x^{(a-1)+1}(1-x)^{(n-a-1)-1}dx$$

$$\int_0^1 x^{a-1}(1-x)^{b-1}dx = \frac{(n-a-1)}{(a-1)+1}\int_0^1 x^{(a-1)+1}(1-x)^{(n-a-1)-1}dx$$

By induction, this process will continue until the exponent of $(1-x)$ goes to zero, which occurs after $(n-a-1)$ times. Thus by induction,

$$\int_0^1 x^{a-1}(1-x)^{b-1}dx = \frac{(n-a-1)!}{((a-1)+1)((a-1)+2)...((a-1)+(n-a-1))}\int_0^1 x^{(a-1)+(n-a-1)}(1-x)^{(n-a-1)-(n-a-1)}dx$$

$$\int_0^1 x^{a-1}(1-x)^{b-1}dx = \frac{(n-a-1)!}{\frac{((a-1)+(n-a-1))!}{(a-1)!}}\int_0^1 x^{n-2}dx = \frac{(n-a-1)!}{\frac{((a-1)+(n-a-1))!}{(a-1)!}}\int_0^1 x^{n-2} = \frac{(n-a-1)!}{\frac{(n-2)!}{(a-1)!}}\int_0^1 x^{n-2}$$

Going back and remembering that $b = n - a$, if b and a are integers (which they are in the identiy on page 33)

$$\int_0^1 x^{a-1}(1-x)^{b-1}dx = \frac{(b-1)!}{\frac{(n-2)!}{(a-1)!}}\int_0^1 x^{n-2} = \frac{\Gamma(b)}{\frac{(n-2)!}{\Gamma(a)}}\int_0^1 x^{n-2}$$

Since

$$(\theta - 1)! = \Gamma(\theta), \forall \theta \in \mathbb{N}$$

Thus,

$$\int_0^1 x^{a-1}(1-x)^{b-1}dx = \frac{\Gamma(b)}{\frac{(n-2)!}{\Gamma(a)}} \int_0^1 x^{n-2} = \frac{\Gamma(a)\Gamma(b)}{(n-2)!} \int_0^1 x^{n-2}$$

$$\int_0^1 x^{a-1}(1-x)^{b-1}dx = \frac{\Gamma(a)\Gamma(b)}{(n-2)!} \int_0^1 x^{n-2} = \frac{\Gamma(a)\Gamma(b)}{(n-2)!} \left[\frac{x^{n-1}}{n-1}\right]_0^1$$

$$\frac{\Gamma(a)\Gamma(b)}{(n-2)!} \left[\frac{x^{n-1}}{n-1}\right]_0^1 = \frac{\Gamma(a)\Gamma(b)}{(n-1)!} \left[x^{n-1}\right]_0^1 = \frac{\Gamma(a)\Gamma(b)}{(n-1)!}$$

And since $n = a + b$

$$\int_0^1 x^{a-1}(1-x)^{b-1}dx = \frac{\Gamma(a)\Gamma(b)}{(n-1)!} = \frac{\Gamma(a)\Gamma(b)}{(a+b-1)!} = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

$$\#QED$$

If you didn't like my assumption that $a, b \in \mathbb{N}$, here is an alternative proof:

$$\Gamma(a)\Gamma(b) = \int_0^\infty x^{a-1}e^{-x}dx \int_0^\infty y^{b-1}e^{-y}dy = \int_0^\infty \int_0^\infty x^{a-1}e^{-x}y^{b-1}e^{-y}dxdy$$

if we let $x = yt$, $dx = ydt$

$$= \int_0^\infty \int_0^\infty (yt)^{a-1}e^{-yt}y^{b-1}e^{-y}ydtdy = \int_0^\infty \int_0^\infty t^{a-1}e^{-y(1+t)}y^{a+b-1}dtdy$$

letting $s = (1+t)v$, $ds = (1+t)dv$,

$$\int_0^\infty \int_0^\infty t^{a-1}e^{-y(1+t)}y^{a+b-1}dtdy = \int_0^\infty \int_0^\infty t^{a-1}e^{-s}\left(\frac{s}{1+t}\right)^{a+b-1}\frac{1}{1+t}dtds$$

$$= \int_0^\infty t^{a-1}\frac{1}{1+t} \int_0^\infty e^{-s}\left(\frac{s}{1+t}\right)^{a+b-1}dtds = \int_0^\infty t^{a-1}(1+t)^{-(a+b)}dt \int_0^\infty (s)^{a+b-1}e^{-s}ds$$

which, by definition of $\Gamma(a+b)$,

$$\int_0^\infty t^{a-1}(1+t)^{-(a+b)}dt \int_0^\infty (s)^{a+b-1}e^{-s}ds = \Gamma(a+b) \int_0^\infty t^{a-1}(1+t)^{-(a+b)}dt$$

now, let $t = \frac{q}{1-q}$, $dt = (\frac{1}{1-q})^2$, with new bounds of 0 to 1 by substituting in q for t, so that

$$\Gamma(a+b)\int_0^\infty t^{a-1}(1+t)^{-(a+b)}dt = \Gamma(a+b)\int_0^1 \left(\frac{q}{1-q}\right)^{a-1}\left(1+\frac{q}{1-q}\right)^{-(a+b)}\left(\frac{1}{1-q}\right)^2 dq$$

$$= \Gamma(a+b)\int_0^1 \left(\frac{q}{1-q}\right)^{a-1}\left(\frac{1}{1-q}\right)^{-(a+b)}\left(\frac{1}{1-q}\right)^2 dq$$

$$= \Gamma(a+b)\int_0^1 q^{a-1}\left(\frac{1}{1-q}\right)^{a-1-(a+b)+2} dq = \Gamma(a+b)\int_0^1 q^{a-1}\left(\frac{1}{1-q}\right)^{1-b} dq$$

$$= \Gamma(a+b)\int_0^1 q^{a-1}(1-q)^{b-1}dq = \Gamma(a+b)\text{beta}(a,b)$$

$$\implies$$

$$\Gamma(a)\Gamma(b) = \Gamma(a+b)\text{beta}(a,b)$$

$$\implies$$

$$\text{beta}(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$