S721 HW4

John Koo

Problem 1.39

Part a

If A and B are mutually exclusive, then $A \cap B = \emptyset \implies P(A \cap B) = 0$.

Assume that A and B are independent. Then $P(A \cap B) = P(A)P(B)$. But the left hand side is equal to 0, which means $P(A) = 0$ or $P(B) = 0$, which is a contradiction.

Part b

If A and B are independent, then $P(A \cap B) = P(A)P(B)$.

Assume that A and B are mutually exclusive. Then $P(A \cap B) = 0$. But this is equal to $P(A)P(B)$. Therefore, $P(A) = 0$ or $P(B) = 0$, which is a contradiction.

Problem 1.40

Part b

$P(A^c \cap B) = P(A^c \mid B)P(B) = (1 - P(A \mid B))P(B)$. Since A and B are independent, $P(A \mid B) = P(A)$. Therefore, $(1 - P(A \mid B)P(B) = (1 - P(A))P(B) = P(A^c)P(B)$.

Part c

$P(A^c \cap B^c) = P(B^c \mid A^c)P(A^c) = (1 - P(B \mid A^c))P(A^c) = (1 - P(B))P(A^c)$ (from part (b))
$= P(B^c)P(A^c)$

Problem 2.1.b

$Y = 4X + 3 \implies X = \frac{Y-3}{4}$

Since $X > 0$, $Y > 3$.

$f_Y(y) = f_X(x(y))\left|\frac{dx}{dy}\right|$
$= 7\exp\big(-\frac{7(y-3)}{4}\big)\frac{1}{4}$
$= \frac{7}{4}\exp\big(-\frac{7}{4}(y-3)\big)$

Then if we integrate,
$\int_{y=3}^{\infty} f_Y(y)dy$
$= \int_3^{\infty} \frac{7}{4}\exp\big(-\frac{7}{4}(y-3)\big)dy$
$= -\exp\big(-\frac{7}{4}(y-3)\big)\Big|_3^{\infty}$
$= -\exp(-\infty) + \exp(0) = 1$

Problem 2.2.b

$Y = -\log X \implies X = \exp(-Y) \implies \frac{dx}{dy} = -\exp(-Y)$

Since $0 < x < 1$, $0 < y < \infty$.

$f_Y(y) = f_X(x(y))\left|\frac{dx}{dy}\right|$
$= \frac{(n+m+1)!}{n!m!}\exp(-ny)(1-\exp(-y))^m\left|-\exp(-y)\right|$
$= \frac{(n+m+1)!}{n!m!}\exp\big(-y(n+1)\big)\big(1-\exp(-y)\big)^m$

Problem 2.5

Using Figure 2.1.1, we can divide the domain into four regions, $(0, \pi/2]$, $(\pi/2, \pi]$, $(\pi, 3\pi/2]$, and $(3\pi/2, 2\pi)$. In the first and third regions, $x = \arcsin(\sqrt{y})$, in the second region, $x = \pi - \arcsin(-\sqrt{y})$, and in the fourth region, $x = 2\pi - \arcsin(-\sqrt{y})$. In any case, $\left|\frac{dx}{dy}\right| = \frac{1}{2\sqrt{y(1-y)}}$.

Then $f_Y(y) = f_X(x(y))\left|\frac{dx}{dy}\right| \times 4$ (four regions)
$= \frac{1}{2\pi}\frac{1}{2\sqrt{y(1-y)}} \times 4$
$= \frac{1}{\pi\sqrt{y(1-y)}}$

Not from text

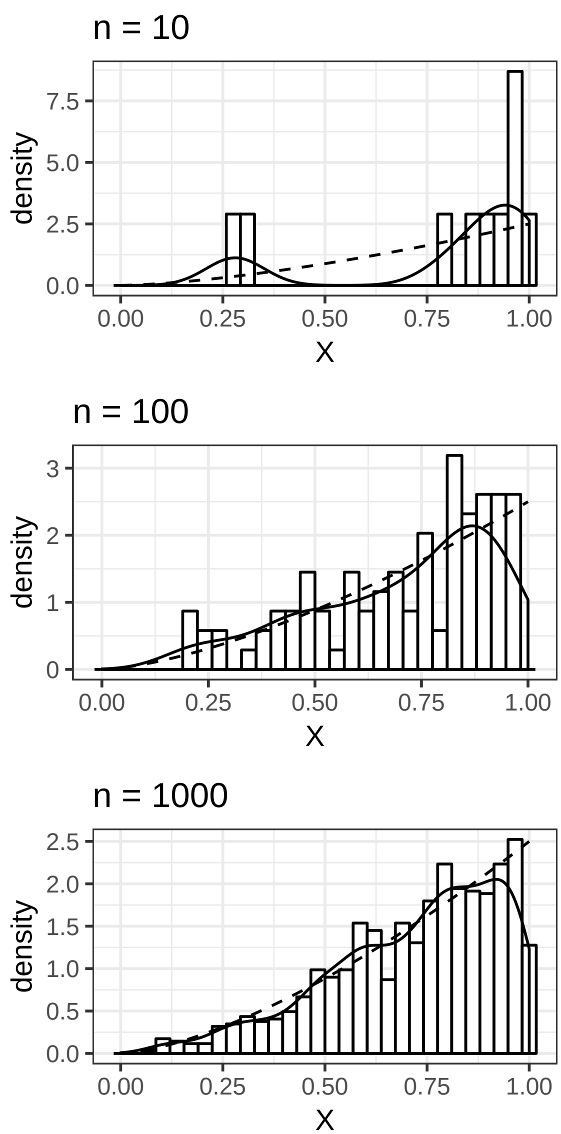
$F(x \mid \theta) = \int_0^x \frac{1}{B(\theta,1)} t^{\theta-1} dt$
$= \frac{x^\theta}{\theta B(\theta,1)}$

Then $F^{-1}(y) = \big(\theta B(\theta,1)y\big)^{1/\theta}$, which we will use to draw our samples.

Note that $B(\theta,1) = \frac{\Gamma(\theta)\Gamma(1)}{\Gamma(\theta+1)} = \frac{\Gamma(\theta)}{\Gamma(\theta)\theta} = \frac{1}{\theta}$. Then our pdf is simply $f(x \mid \theta) = \theta x^{\theta-1}$ and our cdf is simply $F(x \mid \theta) = x^\theta$ (so $F^{-1}(y) = y^{1/\theta}$).

```
library(ggplot2)
import::from(magrittr, `%>%`)

theme_set(theme_bw())

# as specified by the problem
theta <- 2.5
n.vector <- c(10, 100, 1000)

draw.beta <- function(n, theta) {
  # draw from the beta distribution with parameters theta and 1
  Y <- runif(n)
  Y ^ (1 / theta)
}

# ground truth data
dx <- 1e-3
x <- seq(0, 1, by = dx)
y <- theta * x ^ (theta - 1)
```

```r
# construct plots for each n
plots <- lapply(n.vector, function(n) {
  X <- draw.beta(n, theta)
  ggplot() +
    geom_histogram(aes(x = X, y = ..density..),
                   colour = 'black', fill = 'white') +
    geom_density(aes(x = X)) +
    geom_line(aes(x = x, y = y), linetype = 2) +
    labs(title = paste('n =', n)) %>%
    return()
})

do.call(gridExtra::grid.arrange, plots)
```



In the above plots, the solid line is the estimated density from the sample and the dashed line is the ground truth. We can see that as n grows, we get a better approximation to the ground truth.