STATS 551 HW1

Minxuan Chen

2023/9/17

Problem 1

Since a population is partitioned into disjoint groups, we should have

$$P(H_1) + P(H_2) + P(H_3) = 1$$

We first calculate $P(H_i|E)$ explicitly.

$$P(H_i|E) = \frac{P(E|H_i)P(H_i)}{\sum_{i=1}^{3} P(E|H_i)P(H_i)}$$

from this, we observe that $P(H_i|E)$ all have the same denominator, so the comparisions among them only depend on the numerator.

For $P(E|H_i)$, we have

$$P(E|H_1) = 0.1 < P(E|H_2) = 0.3 < P(E|H_3) = 0.5$$

therefore

(a)

$$P(H_1) = 0.1, \quad P(H_2) = 0.8, \quad P(H_3) = 0.1$$

(b)

$$P(H_1) = 0.4, \quad P(H_2) = 0.1, \quad P(H_3) = 0.5$$

(c)

$$P(H_1) = 0.8, \quad P(H_2) = 0.1, \quad P(H_3) = 0.1$$

(d)

$$P(H_1) = 0.7, \quad P(H_2) = 0.2, \quad P(H_3) = 0.1$$

(e)

$$P(H_1) = 0.6, \quad P(H_2) = 0.3, \quad P(H_3) = 0.1$$

Check

```r
pE_H <- c(0.1,0.3,0.5)
pH <- matrix(c(0.1,0.4,0.8,0.7,0.6,
               0.8,0.1,0.1,0.2,0.3,
               0.1,0.5,0.1,0.1,0.1),ncol = 3)

pH_E <- t(pE_H*t(pH))/c(pH%*%pE_H)

#order
t(apply(pH_E, 1, order))
```

```
##      [,1] [,2] [,3]
## [1,]    1    3    2
## [2,]    2    1    3
## [3,]    2    3    1
## [4,]    3    2    1
## [5,]    3    1    2
```

Problem 2

Let L_1, L_2, L_3 denote the car is located behind Door 1,2,3, resp. Let H_1, H_2, H_3 denote the host open Door 1,2,3, resp.

Before the game,

$$P(L_1) = P(L_2) = P(L_3) = \frac{1}{3}$$

The probability of winning (if switching) is

$$\begin{aligned}
P(\text{win}|H_3) &= P(L_2|H_3) \\
&= \frac{P(H_3|L_2)P(L_2)}{P(H_3|L_1)P(L_1) + P(H_3|L_2)P(L_2) + P(H_3|L_3)P(L_3)} \\
&= \frac{P(H_3|L_2)}{P(H_3|L_1) + P(H_3|L_2) + P(H_3|L_3)}
\end{aligned}$$

For these conditional probabilities, $H_3|L_3$ is impossible, so $P(H_3|L_3) = 0$. For $H_3|L_2$, the host has not choice rather opening Door 3, so $P(H_3|L_2) = 1$.
For $H_3|L_1$, since the car is in neither Door 2 nor 3, the host can random choose from the two doors. Note that the prob host open Door 2 or 3 depends on his location.

(1) Suppose the host is closer to Door 2, then $P(H_3|L_1) = 1 - \gamma \in (0, 1/2]$. So

$$P(\text{win}|H_3) = \frac{1}{1-\gamma+1} = \frac{1}{2-\gamma} \in [\frac{2}{3}, 1)$$

Since $P(\text{win}|H_3) \geq \frac{2}{3} > \frac{1}{2}$, the probability of winning is higher by switching.

(2) Suppose the host is closer to Door 3, then $P(H_3|L_1) = \gamma \in [1/2, 1)$. So

$$P(\text{win}|H_3) = \frac{1}{\gamma + 1} \in (\frac{1}{2}, \frac{2}{3}]$$

Since $P(\text{win}|H_3) > \frac{1}{2}$, the probability of winning is higher by switching.

By (1),(2), no matter where the host is, under $\gamma \in [1/2, 1)$, the player will have a better chance of winning by always switching the door

Problem 3

It's easy to use a Bayesian Network to represent their relationships. We use the following information to construct this DAG.

1. X Poisson
2. Y use PMF, use population/total population ECDF
3. Z Bernoulli.
4. U Exponential./ truncated normal
5. V truncated normal
6. W ??
7. Country (Y) influences age (X) structure.
8. Country (Y) and age (X) influence whether the person attended a picnic (Z), since country reflects cultural and different age has different opinion and tendency for a picnic.

9. Age (X) influences coffee consumed (U) and screen time spent (V). For example, young age needs working, so maybe more coffee and screen time.
10. Screen time (V) influences coffee consumed(U).
11. Country (Y) influences passport photo (W).
12. All the term "influence" above means Markovian parents of the variable.
13. All distribution above also means conditional distribution.

DAG:

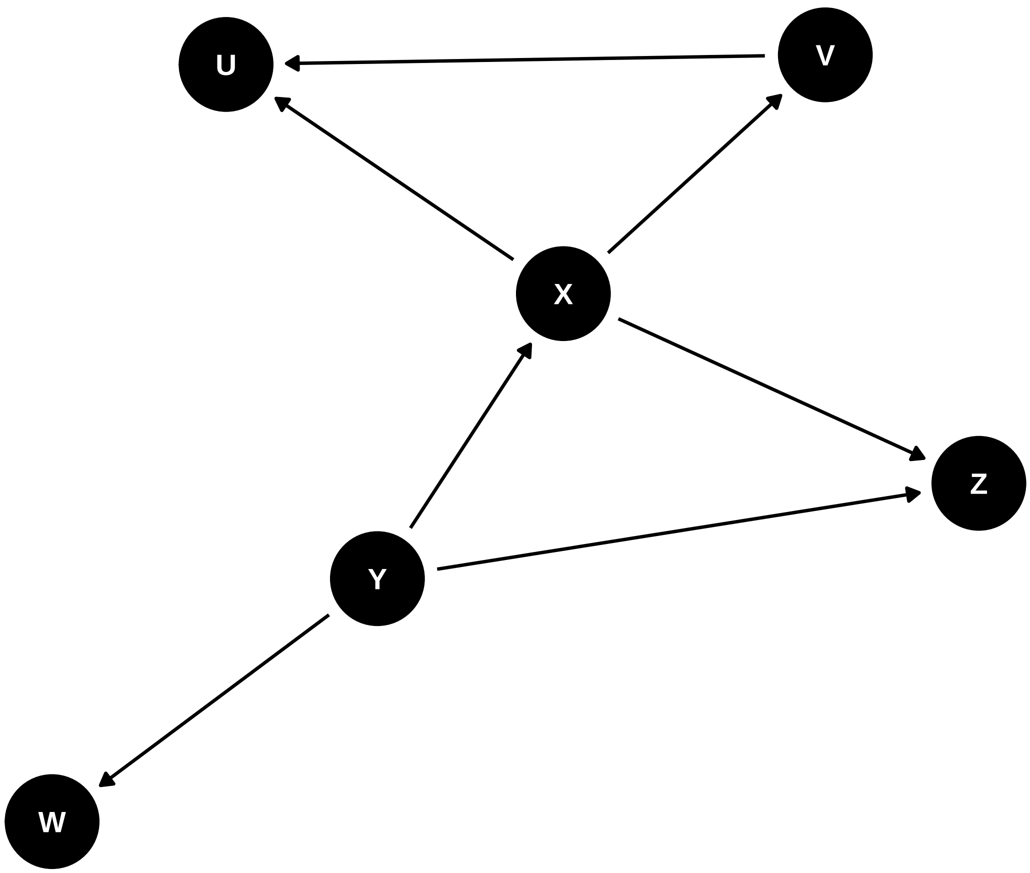
```
Sys.setenv(LANGUAGE = "en")
library(dagitty)
library(ggdag)
```

```
##
## Attaching package: 'ggdag'

## The following object is masked from 'package:stats':
##
##     filter
```

```
dag <- dagitty::dagitty("
       dag{
       Y->W
       Y->Z
```

```
        Y->X->U
        Y->X->V->U
        Y->X->Z
        }
         ")
tidy_dag <- tidy_dagitty(dag)
ggdag(tidy_dag)+theme_dag()
```



Since there's no collider or any separate node in this DAG, we conclude that X, Y, Z, U, V, W are not pairwise independently distributed.

As for joint distribution, it's easy to read from the DAG.

$$f(x, y, z, u, v, w) = f(x|y)f(y)f(z|x, y)f(u|x, v)f(v|x)f(w|y)$$

Problem 4

$$\theta \sim \text{Beta}(a, b) : p(\theta) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\theta^{a-1}(1-\theta)^{b-1}$$

Therefore

$$\text{mode}[\theta] = \max_\theta \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\theta^{a-1}(1-\theta)^{b-1}$$

$$\frac{dp(\theta)}{d\theta} = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\left((a-1)\theta^{a-2}(1-\theta)^{b-1} - \theta^{a-1}(b-1)(1-\theta)^{b-2}\right) = 0$$

we have

$$(a-1)\theta^{a-2}(1-\theta)^{b-1} - \theta^{a-1}(b-1)(1-\theta)^{b-2} = 0 \to \theta = \frac{a-1}{a+b-2}$$

Note that if $a > 0, b > 0$, $p(\theta = 0) = p(\theta = 1) = 0$, since $p(\theta) \geq 0 \, \forall\theta$ and $p(\theta)$ is continuous, The mode we get above can maximize $p(\theta)$.

$$\begin{aligned}
\mathbb{E}(\theta) &= \int_0^1 \theta \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\theta^{a-1}(1-\theta)^{b-1}d\theta \\
&= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\int_0^1 \theta^{a+1-1}(1-\theta)^{b-1}d\theta \\
&= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\frac{\Gamma(a+1)\Gamma(b)}{\Gamma(a+b+1)} \\
&= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\frac{a\Gamma(a)\Gamma(b)}{(a+b)\Gamma(a+b)} \\
&= \frac{a}{a+b} \\
\mathbb{E}\theta^2 &= \int_0^1 \theta^2 \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\theta^{a-1}(1-\theta)^{b-1}d\theta \\
&= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\int_0^1 \theta^{a+2-1}(1-\theta)^{b-1}d\theta \\
&= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\frac{\Gamma(a+2)\Gamma(b)}{\Gamma(a+b+2)} \\
&= \frac{a(a+1)}{(a+b)(a+b+1)}
\end{aligned}$$

thus

$$\text{Var}[\theta] = \mathbb{E}\theta^2 - (\mathbb{E}\theta)^2 = \frac{a(a+1)}{(a+b)(a+b+1)} - \frac{a^2}{(a+b)^2} = \frac{ab}{(a+b)^2(a+b+1)}$$

Problem 5

The parameter we care about is θ, which means the probability of head of the chosen coin. Our prior is $P(\text{heads}|C_1) = 0.6, P(\text{heads}|C_2) = 0.4$, since we choose one coin at random,

$$P(\theta = 0.6) = P(\theta = 0.4) = \frac{1}{2}$$

Use T_1, T_2 to represent the first two spins are tails. This event is data and

$$T_i|\theta \sim \text{Bern}(\theta), \text{i.i.d}$$

Posterior of θ is

$$p(\theta|T_1T_2) = \frac{p(T_1T_2|\theta)p(\theta)}{p(T_1T_2)}$$

we have

$$p(\theta = 0.6|T_1T_2) = \frac{4}{13}, p(\theta = 0.6|T_1T_2) = \frac{9}{13}$$

Let Y denote the number of additional spins,

$$Y|\theta \sim \text{Geom}(\theta)$$

Therefore

$$\begin{aligned}
\mathbb{E}(Y|T_1T_2) &= \mathbb{E}(\mathbb{E}(Y|\theta, T_1T_2)|T_1T_2) \\
&= \mathbb{E}(\mathbb{E}(Y|\theta)|T_1T_2) \\
&= \mathbb{E}\left(\frac{1}{\theta}\middle| T_1T_2\right) \\
&= \frac{4}{13}\cdot\frac{1}{0.6} + \frac{9}{13}\cdot\frac{1}{0.4} \\
&= \frac{175}{78}
\end{aligned}$$