BST 140.651 Problem Set 1

Kai Aragaki

9/10/2020

Problem 1. Show the following

a.$P(\emptyset) = 0$.

$\Omega = \emptyset \cup \Omega$
$P(\Omega) = P(\emptyset \cup \Omega)$
$P(\emptyset \cup \Omega) = P(\emptyset) + P(\Omega)$
So,
$P(\Omega) = P(\emptyset) + P(\Omega)$
Since $P(\Omega) = 1$,
$1 = P(\emptyset) + 1$
And solving for $P(\emptyset)$ we have:
$P(\emptyset) = 0$

Because $\emptyset \cap \Omega = \emptyset$ (ie the two sets are non-overlapping)

■

b. $P(E) = 1 - P(E^c)$.

$1 = P(\Omega)$
$1 = P(E \cup E^c)$
$1 = P(E) + P(E^c)$
Solving for $P(E)$,
$P(E) = 1 - P(E^c)$

Because $E \cap E^c = \emptyset$

■

c. if $A \subset B$, then $P(A) \leq P(B)$

If $A \subset B$,
$B = A \cup (A^c \cap B)$
$P(B) = P(A \cup (A^c \cap B))$
$P(B) = P(A) + P(A^c \cap B))$
$P(B) - P(A) = P(A^c \cap B)$
We know that
$P(A^c \cap B) \geq 0$
because
$0 \leq P(E) \leq 1$
therefore,
$P(B) - P(A) \geq 0$
Solving for $P(B)$ we arrive at
$P(B) \geq P(A)$

Because $A \cap (A^c \cap B) = \emptyset$

■

d. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$A = A \cap \Omega = A \cap (B \cup B^c) = (A \cap B) \cup (A \cap B^c)$
taking the probability of both sides,
$P(A) = P((A \cap B) \cup (A \cap B^c))$
Because $(A \cap B) \cap (A \cap B^c) = \emptyset$,
$P(A) = P(A \cap B) + P(A \cap B^c)$
By the same logic,
$P(B) = P(A \cap B) + P(B \cap A^c)$
We can substitute these into the first equation, giving
$P(A \cap B) + P(A \cap B^c) + P(A \cap B) + P(B \cap A^c) - P(A \cap B)$
Simplifying,
$P(A \cup B) = P(A \cap B^c) + P(A \cap B) + P(B \cap A^c)$
Since we know that the sum of probabilities of two disjoint sets is
equal to probability of their union, and we know $(A \cap B) \cap (A \cap B^c) = \emptyset$
so this is equivalent to
$P((A \cap B^c) \cup (A \cap B)) + P(B \cap A^c)$
and by the distributive property,
$P(A \cap (B \cup B^c)) + P(B \cap A^c)$
Which simplifies to
$P(A) + P(B \cap A^c)$
We know that
$A \cap (B \cap A^c) = \emptyset$
because
$A \cap A^c = \emptyset$
and
$(B \cap A^c) \subset A^c$.
So,
$P(A \cup (B \cap A^c))$
which again by the distributive property,
$P((A \cup B)) \cap (A \cup A^c))$
and,
$P((A \cup B)) \cap \Omega) = P(A \cup B)$ ∎

e. $P(A \cup B) = 1 - P(A^c \cap B^c)$

$1 = P(\Omega) = P((A \cup B) \cup (A \cup B)^c)$
Given the sets are disjoint,
$P((A \cup B) \cup (A \cup B)^c) = P(A \cup B) + P(A \cup B)^c$
By Demorgan's law,
$P(A \cup B) + P(A \cup B)^c = P(A \cup B) + P(A^c \cap B^c)$
Substituting this into the statement we intend to prove,

$P(A \cup B) = P(A \cup B) + P(A^c \cap B^c) - P(A^c \cap B^c)$
$P(A \cup B) = P(A \cup B)$ ∎

f. $P(A \cap B^c) = P(A) - P(A \cap B)$

Rearranging the above, we note
$P(A \cap B^c) + P(A \cap B) = P(A)$
Since $(A \cap B^c) \cap (A \cap B) = \emptyset$,
$P((A \cap B^c) \cup (A \cap B)) = P(A)$
Which by the distributive property is equal to
$P(A \cap (B^c \cup B))) = P(A \cap \Omega) = P(A)$ ∎

g. $P(\cup_{i=1}^n E_i) \leq \Sigma_{i=1}^n P(E_i))$

Proof by induction, starting with the case $n = 2$.
We know that
$P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$ see proof d
And since $P(E_1) + P(E_2) = \Sigma_{i=1}^2 P(E_i)$
by substitution we get
$P(E_1 \cup E_2) = \Sigma_{i=1}^2 P(E_i) - P(E_1 \cap E_2)$
Since the probability of any even must be greater than or equal to
zero, the right hand side of the equation (and by definition the left)
must be less than or equal to $\Sigma_{i=1}^2 P(E_i)$ itself! So,
$P(E_1 \cup E_2) \leq \Sigma_{i=1}^2 P(E_i)$
Since we have proven our statement for n = 2, let us assume it true
for n-1:
$P(\cup_{i=1}^{n-1} E_i) \leq \Sigma_{i=1}^{n-1} P(E_i)$
Then for n,
$P(\cup_{i=1}^n E_i) = P(E_n) \cup P(\cup_{i=1}^{n-1} E_i)$
We notice that this is structurally identical to that which we proved
above, where $P(E_1) = P(E_n)$ and $P(E_2) = P(\cup_{i=1}^{n-1} E_i)$. Following the
same proof we find
$P(\cup_{i=1}^n E_i) \leq \Sigma_{i=1}^n P(E_i))$ ∎

h. $P(\cup_{i=1}^n E_i) \geq max_i P(E_i)$

For proof by induction, begin with n = 2:
$P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$ proof d
If we suppose $P(E_1) = max_i P(E_i)$, then this is
$P(E_1) + P(E_2) - P(E_1 \cap E_2) \geq P(E_1)$

Which we can simplify to

$P(E_2) \geq P(E_1 \cap E_2)$

which we know to be true   proof c

We now consider the case for n-1 to be true:

$P(\cup_{i=1}^{n-1} E_i) \geq max_i P(E_i)$

and evaluate the truth of n:

$P(E_n) \cup P(\cup_{i=1}^{n-1} E_i) \geq max_i P(E_i)$

Like proof g, we can state that

$P(E_n)$ is analogous to $P(E_1)$

and

$P(\cup_{i=1}^{n-1} E_i)$ is analgous to $P(E_2)$,

at which point this is a statement we have already proved   ∎

Problem 2

Cryptosporidium is a pathogen that can cause gastrointestinal illness with diarrhea; infections can lead to death in individuals with a weakened immune system. During a recent outbreak of cryptosporidiosis in 21% of two parent families at least one of the parents has contracted the disease. In 9% of the families the father has contracted cryptosporidiosis while in 5% of the families both the mother and father have contracted cryptosporidiosis.

a. What event does the probability one minus the probability that both have contracted cryptosporidiosis represent?

_____

The probability that one or fewer parents in a two parent family contracted the disease

b. What's the probability that either the mother or the father has contracted cryptosporidiosis?

_____

21%, if by this we include the possibility of the mother AND father having contracted cryptosporidiosis If we do not include the possibility of the mother AND father, it is $P(father \cup mother) - P(father \cap mother)$ which is 21%-5%, or

16%

c. What's the probability that the mother has contracted cryptosporidiosis but the father has not?

_____

We can solve for $P(mother)$ first with knowledge that
$$P(father \cup mother) = P(father) + P(mother) - P(father \cap mother)$$
Substituting in our given values,
$$21\% = 9\% + P(mother) - 5\%$$
which solving for $P(mother)$ gives 17%. Subtracting the intersect
gives

12%

d. *What's the probability that the mother has contracted cryptosporidiosis?*

_____

This was calculated above to be
17%

e. *What's the probability that neither the mother nor the father has contracted cryptosporidiosis?*

_____

100% - 21% = **79%**

f. *What's the probability that the mother has contracted cryptosporidiosis but the father has not?*

Identical to problem c

Problem 3

_____

A valid pmf follows two rules: $0 \geq p(h(x)) \geq 1$ for all values and $\Sigma_1^I p(x) = 1$.

To the first point, all input values are positive, so all probabilities are positive by this function. Additionally, $h(x) \leq \Sigma_{i=1}^I h(i)$ since $\Sigma_{i=1}^I h(i)$ is a sum of non-zero values (given) from $h(x)$, so
$$\Sigma_{i=1}^I h(i) \geq h(x) \implies h(x)/\Sigma_{i=1}^I h(i) \leq 1 \implies p(x) \leq 1$$
To the second point,
$$\Sigma_1^I p(x) = \Sigma_1^I (h(x)/\Sigma_1^I h(x)) = \Sigma_1^I h(x)/\Sigma_1^I h(x) = 1 \qquad \blacksquare$$

Problem 4

_____

A valid pdf follows two rules: $0 \geq f(h(x)) \geq 1$ across the function, and $\int f(x)dx = 1$ for all values.

Since we know that $h > 0$, all possible values of the function $f(x)$ can also only be positive.

We can show by similar logic above (albeit on an infinitesimal level) that $h(x) \leq c$ since c is a sum of non-zero (given) values from $h(x)$. Therefore $h(x)/c \leq 1$.

To the second point,

$\int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^{\infty} h(x)dx/c$
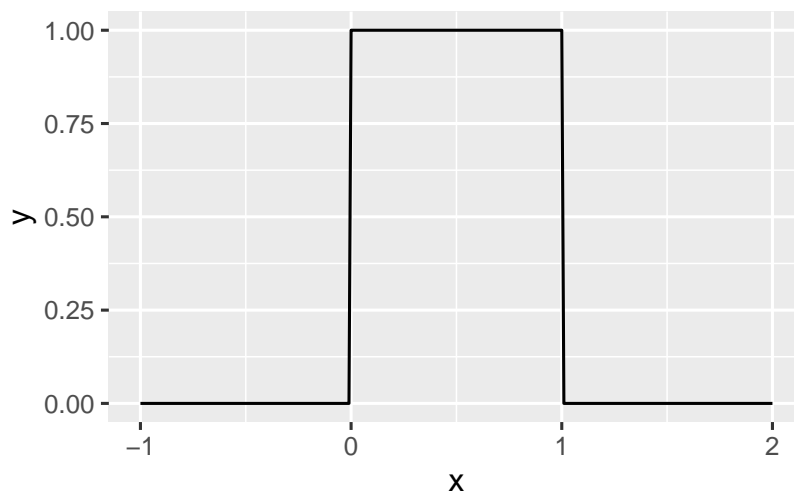
but we were told

$c = \int_{-\infty}^{\infty} h(x)dx$

so

$\int_{-\infty}^{\infty} f(x)dx = c/c = 1$ ∎

Problem 5

_____

a.

_____

```
dat <- tibble(x = seq(-1, 2, by = 0.01), y = dunif(x))
ggplot(dat, aes(x = x, y = y)) + geom_line()
```


Since we know that the area under the curve of a PDF must be equal to 1, and the distribution goes from 0 to 1, k must equal 1.

b

_____

Mathematically we can represent this as

$\int_{.7}^{.1} f(x) = .6$

In the context of the problem, we can think about this as "what is the probability that there exists an individual within this population that has proportion of their skin covered in freckles somewhere between 0.1 and 0.7?

c

```
punif(0.7) - punif(0.1)

## [1] 0.6
```

We can see that for this particular distribution, the probability that $a < X < b$ given $0 < a < b < 1$ is $b - a$

d

---

pdf

e

The median (or 50%) is the probability (or probability density) at the point where half the values are greater than the given value, and half are lower. In the context of the uniform distribution, the median probability is 0.5. In the context of our problem, this states that half the people in the population have greater than half their skin covered in freckles, while the other half of the population have less than half their skin covered in freckles.

f

Here the 95% is 0.95 (lower tail) - that is, in the context of this problem, 95% of the population will have a smaller proportion of their skin befreckled

g

This seems unlikely to me. In the real world there are given sub-populations that tend to have similar amounts of clusters (tending more towards a collection of gaussians vs a completely smooth uniform

distribution). Intuitively it also seems strange that one could have 0 freckles just as probably as one could have 1 freckle - it seems more likely that if one has one freckle, there are probably more.

Problem 6

_____

a

_____

A random variable U is passed to a function $F^{-1}(U)$ which we will define as a new random variable, X.

Our intent is to find that $CDF(X) = F$

We define a CDF as the probability of obtaining a value of or less than that value. That is,

$CDF(X) = Pr(X \leq x)$

or, using our definition of X above,

$Pr(X \leq x) = Pr(F^{-1}(U) \leq x)$

We were told that F was strictly increasing, and is therefore invertable. We take F of both sides of the inequality to get

$Pr(F^{-1}(U) \leq x) = Pr(F(F^{-1}(U) \leq F(x))$

and use the knowledge that for invertable functions,

$F(F^{-1}(Y)) = Y$

so

$Pr(F(F^{-1}(U) \leq F(x)) = Pr(U \leq F(x))$

The CDF can also be rewritten as

$\int_{lowerbound}^{y} pdf(Y)dy$

Where Y is a random variable. In this instance we know that the pdf of U is a uniform distribution, which for $[0, 1]$ is 1. Additionally, in this instance our upper bound is $F(x)$, so

$\int_{0}^{F(x)} 1dx = F(x)$ ∎

b

_____

Since we know that values of $F^{-1}(U)$ have a cdf equal to F, we can generate n simulated points by doing

```
set.seed(123)
n <- 100
points <- runif(100, min = 0, max = 1)
```

then we can feed the `points` argument into our function F to get a set of 100 points with a cdf approximating F.

c

```
set.seed(123)
n <- 100
points <- runif(100, min = 0, max = 1)
out <- dnorm(points, mean = 0.5, sd = 0.25)
plot(points, out)
```