Conditionality and stopping times in probability

Mark Osegard, Ben Speidel, Megan Silberhorn, and Dickens Nyabuti
Conditional Expectation
Conditional Probability

Discrete: Conditional Probability Mass Function

\[ P\{X = x \mid Y = y\} = \frac{P\{X = x, Y = y\}}{P\{Y = y\}} \]

Continuous: Conditional Probability Density Function

\[ f_{X \mid Y}(x \mid y) := \frac{f(x, y)}{f_X(y)} \]
Conditional Expectation

Discrete: \[ E[X \mid Y = y] = \sum_x xP\{X = x \mid Y = y\} \]

Continuous: \[ E[X \mid Y = y] = \int_{-\infty}^{+\infty} xf_{X \mid Y}(x, y) dx \]
Note:

\[ y \mapsto E[X \mid Y = y] \text{ is a function of } y. \text{ We write this as } E[X \mid Y] \]

i.e. \[ E[X \mid Y](y) = E[X \mid Y = y] \]

(Conditional Expectation Function)
Theorem:

\[ E[X] = E[E[X \mid Y]] \]

Clearly, when \( Y \) is discrete,

\[ = \sum_{y} E[X \mid Y = y] P\{Y = y\} \]

When \( Y \) is continuous,

\[ = \int_{-\infty}^{+\infty} E[X \mid Y = y] f_{Y}(y) dy \]
Proof: Continuous Case

Recall, if $X,Y$ are jointly continuous with joint pdf $f(x, y)$

Define: $f_{X | Y}(X | Y) = \frac{f(x, y)}{f_Y(y)}$

and $E[X | Y = y] = \int_{-\infty}^{+\infty} xf_{X | Y}(X | Y)dx$
Note:

\[
\int_{-\infty}^{+\infty} E[X \mid Y = y] f_Y(y) dy = \int_{-\infty}^{+\infty} \left( \int_{-\infty}^{+\infty} x f_{X \mid Y}(x \mid y) dx \right) f_Y(y) dy
\]

\[
= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x f_{X \mid Y}(x \mid y) f_Y(y) dx dy
\]

\[
= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x \left( \frac{f(x, y)}{f_Y(y)} \right) f_Y(y) dx dy
\]
Continuous Case Cont.

\[
= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xf(x, y) \, dx \, dy = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xf(x, y) \, dy \, dx
\]

(Fubini’s Theorem)

\[
= \int_{-\infty}^{+\infty} x \int_{-\infty}^{+\infty} f(x, y) \, dy \, dx = \int_{-\infty}^{+\infty} f(x, y) \, dy
\]
So,

\[ \int_{-\infty}^{+\infty} x f_X(x) \, dx = E[X] \]

Therefore, concluding

\[ E[X] = E[E[X \mid Y]] \]
Summary:

When $Y$ is discrete,

$$= \sum_y E[X \mid Y = y]P\{Y = y\}$$

When $Y$ is continuous,

$$= \int_{-\infty}^{+\infty} E[X \mid Y = y]f_Y(y)dy$$
Conditional Variance
Definition

\[ Var(X \mid Y) = E[(X - E[X \mid Y])^2 \mid Y] \]

\[ = E[X^2 \mid Y] - (E[X \mid Y])^2 \]

\[ Var(X) = E[Var(X \mid Y)] + Var(E[X \mid Y]) \]
Proof

Using

\[ E[X] = E[E[X | Y]] \]

since

\[ \text{Var}(X | Y) = E[X^2 | Y] - (E[X | Y])^2 \]

taking expectations of both sides

\[ E[\text{Var}(X | Y)] = E[E[X^2 | Y] - (E[X | Y])^2] \]

\[ = E[E[X^2 | Y]] - E[(E[X | Y])^2] \]

\[ = E[X^2] - E[(E[X | Y])^2] \]
Note as well …

\[
\text{Var}(E[X \mid Y]) = E\left[\left(E[X \mid Y]^2\right)\right] - (E[E[X \mid Y]])^2
\]

\[
= E\left[\left(E[X \mid Y]^2\right)\right] - (E[X])^2
\]
...adding

\[ E[\text{Var}(X \mid Y)] + \text{Var}(E[X \mid Y]) = E[X^2] - (E[X])^2 \]

\[ = \text{Var}(X) \]

Thus we've shown that

\[ \text{Var}(X) = E[\text{Var}(X \mid Y)] + \text{Var}(E[X \mid Y]) \nabla \]
Stopping times
Stopping Times

- Definition
- Application to Probability
- Applications of Stopping Times to other formulas
Stopping Times

- Basic Definition:
  A Stopping Time for a process does exactly that, it tells the process when to stop.

  Example: while ( x != 4 )
  {
    ... 
  }

  The stopping time for this code fragment would be the instance where x does equal 4.
Define:

Suppose we have a sequence of Random Variables (all independent of each other)
Our sequence then would be:

\[ X_1, X_2, X_3, \ldots \]
From our previous slide we have the sequence:

\[ X_1, X_2, X_3, \ldots \]

A discrete Random Variable \( N \) is a stopping time for this sequence if:

\[ \{ N = n \} \]

Where \( n \) is independent of all following items in the sequence

\[ X_{n+1}, X_{n+2}, \ldots \]
Summarizing the idea of stopping times with Random Variables we see that the decision made to stop the sequence at Random Variable $N$ depends solely on the values of the sequence

$$X_1, X_2, \ldots, X_n$$

Because of this, we then can see that $N$ is independent of all remaining values $X_m, m > n$
Applications of Stopping Times

Does Stopping Times affect expectation? No!

Consider this statement:

\[ S = \sum_{i=1}^{N} X_i \quad E[S] = E[N]E[X] \]

This formula, the formula used for Conditional Expectation does remain unchanged.
Applying Stopping Times

For an example of how to use stopping times to solve a problem, we will now introduce to you Wald’s Equation...

\[
E\left[ \sum_{i=1}^{N} X_i \right] = E[N]E[X]
\]
Wald’s Equation
If \( \{X_1, X_2, X_3, \ldots \} \) are independent identically distributed (iid) random variables having a finite expectation \( E[X] \), and \( N \) is a stopping time for the sequence having finite expectation \( E[N] \), then:

\[
E \left[ \sum_{i=1}^{N} X_i \right] = E[N]E[X]
\]
Wald’s Proof

Let $N_1 = N$ represent the stopping time for the sequence

$$\{ X_1, X_2, \ldots, X_{N_1} \}$$

Let $N_2 =$ the stopping time for the sequence

$$\{ X_{(N_1+1)}, X_{(N_1+2)}, \ldots, X_{(N_1+N_2)} \}$$

Let $N_3 =$ the stopping time for the sequence

$$\{ X_{(N_1+N_2+1)}, X_{(N_1+N_2+2)}, \ldots, X_{(N_1+N_2+N_3)} \}$$
We can now define the sequence of stopping times as

\[ \{N_i\} \quad i \geq 1 \]

where \( \{N_i\} \) clearly represents,

\[ \{N_1, N_2, N_3, \ldots\} \]

and see the sequence is iid
Wald’s Proof...

If we define a sequence \( \{S_i\} \) as,

where,

\[
\{S_i\} = \{s_1 + s_2 + \ldots + s_m\}
\]

\[
\{S_i\} = \begin{cases} 
S_1 = \sum_{i=1}^{N_1} X_i, & S_2 = \sum_{i=N_1+1}^{N_1+N_2} X_i, \ldots, S_m = \sum_{i=N_1+N_2+\ldots+N_{m-1}}^{N_1+N_2+\ldots+N_m} X_i 
\end{cases}
\]

Note: \( \{S_i\} \) are iid
Wald’s Proof...

\( \{S_i\} = \{S_1 + S_2 + \ldots + S_m\} \) will consist of

\( \{N_1 + N_2 + \ldots + N_m\} \) which are iid because the \( X_i \)'s are.

with common mean \( E[X_i] = E[X] \)
Wald’s Proof...

By the Strong Law of Large Numbers,

\[ \lim_{m \to \infty} \frac{\{S_1 + S_2 + \ldots + S_m\}}{\{N_1 + N_2 + \ldots + N_m\}} = E[X] \]
Wald’s Proof...

Also

\[
E[X] \quad \frac{\{S_1 + S_2 + \ldots + S_m\}}{m} \quad \frac{\{N_1 + N_2 + \ldots + N_m\}}{m}
\]

\[
\downarrow \quad \text{let } m \to \infty \quad \downarrow
\]

\[
E[S] \quad E[N]
\]
Concluding

So as we let \( m \rightarrow \infty \)

\[
E[X] = \frac{E[S]}{E[N]}
\]

Which can be manipulated into our preposition:

\[
E[S] = E \left[ \sum_{i=1}^{N} X_i \right] = E[N]E[X]
\]
Miners Problem

Sample Conditional and Stopping times in probability problem
A miner is trapped in a room containing three doors. Door one leads to a tunnel that returns to the same room after 4 days; door two leads to a tunnel that returns to the same room after 7 days; door three leads to freedom after a 3 day journey. If the miner is at all times equally likely to choose any of the doors, find the expected value and the variance of the time it takes the miner to become free.
Expected Value

Using Wald’s Equation:

\[ X_i \in \{4,7,3\} \]
\[ N = \text{the stopping time} \]
\[ N = \min\{i \mid X_i = 3\} \]
Continue ....

$N$ is a geometric distribution with a Probability $\frac{1}{3}$, selecting door 3 denotes success once this event occurs, its trial Number will be set to $N$ to denote the stopping time.
Continue 

Recall:

The expected value of a Geometric distribution is $\frac{1}{p}$

$\therefore E[N] = \frac{1}{\frac{1}{3}} = 3$
Expected value Conclusion

Substituting using Wald's Equation,

\[ E[X] = E[N]E[X_i] \]

we get

\[ E[X] = 3 \times \frac{14}{3} = 14 \text{ days} \]

Thus on average, the miner is expected to attain freedom in 14 days
Variance

using the equation

\[ \text{Var}(X) = E[\text{Var}(X \mid N)] + \text{Var}(E[X \mid N]) \]

Solution:

\[
E[X \mid N = n] = E \left[ \sum_{i=1}^{n} X_i \mid N = n \right]
\]

\[
= \sum_{i=1}^{n} E[X_i \mid N = n]
\]
Continue.....

\[ E[X_i \mid N = n] = \sum xP\{X_i = x \mid N = n\} \]
\[ = 4P\{X_i = 4 \mid N = n\} + 7P\{X_i = 7 \mid N = n\} \]
\[ + 3P\{X_i = 3 \mid N = n\} \]

\( \therefore E[X_i \mid N = n] = \begin{cases} 
\frac{11}{2} & i < n \\
3 & i = n 
\end{cases} \)


\[
\sum_{i=1}^{n} E[X_i \mid N = n] = \sum_{i=1}^{n-1} E[X_i \mid N = n] + E[X_i \mid N = n]
\]

\[
= \frac{11}{2} (n-1) + 3
\]

\[
= \frac{11}{2} n - \frac{5}{2}
\]

\[
\therefore E[X \mid N] = \frac{11}{2} N - \frac{5}{2}
\]
Continue.....

using

\[ \text{Var} \left( aX + b \right) = a^2 \text{Var} \left( X \right) \]

then

\[ \text{Var} \left( E \left[ X \mid N \right] \right) = \text{Var} \left( \frac{11}{2} N - \frac{5}{2} \right) \]

\[ = \frac{121}{4} \text{Var} \left( N \right) \]

\[ \text{Var} \left( N \right) = \frac{1 - P}{P^2} = \frac{1 - \frac{1}{3}}{\left( \frac{1}{3} \right)^2} = 6 \]
Thus far

\[ Var(X) = E[Var(X \mid N)] + Var(E[X \mid N]) \]

\[ \frac{363}{2} \]
Continue…..

\[ \text{Var}(X \mid N) = \text{Var}\left(\sum_{i=1}^{n} X_i \mid N = n\right) \]

\[ = \sum_{i=1}^{n} \text{Var}(X_i \mid N = n) \]

\[ \text{Var}(X_i \mid N = n) = E[X_i^2 \mid N = n] - (E[X_i \mid N = n])^2 \]

\[ E[X_i^2 \mid N = n] = \sum x^2 P\{X_i = x \mid N = n\} \]
Continue.....

\[= 4^2 P\{X_i = 4 \mid N = n\} + 7^2 P\{X_i = 7 \mid N = n\} + 3^2 P\{X_i = 3 \mid N = n\}\]

\[E[X_i^2 \mid N = n] = \begin{cases} 
\frac{65}{2} & i < n \\
9 & i = n 
\end{cases}\]

\[Var(X_i \mid N = n) = \begin{cases} 
\frac{65}{2} - \frac{121}{4} = \frac{9}{4} & i < n \\
9 - 9 = 0 & i = n 
\end{cases}\]
$Var(X \mid N = n) = \sum_{i=1}^{n-1} Var(X_i \mid N = n) + Var(X_n \mid N = n)$

$= \frac{9}{4} (N - 1)$

$E[Var(X \mid N)] = E\left[\frac{9}{4} (N - 1)\right]$  

$= \frac{9}{4} (E[N] - E[1]) = \frac{9}{4} (3 - 1) = \frac{9}{2}$
In conclusion

\[ \text{Var}(X) = E[\text{Var}(X \mid N)] + \text{Var}(E[X \mid N]) \]

\[ \frac{9}{2} + \frac{363}{2} = \frac{372}{2} = 186 \]

Hence
Thanks to:

Dr. Steve Deckelman
Dr. Ayub Hossain

Who helped make this a success!!!!!!!!!!!!!!!!!!!!!!!!!!!