# Time Series and Stochastic Processes

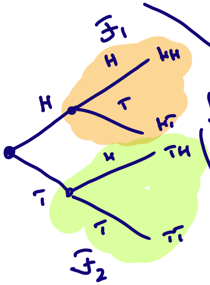
Peter Lukianchenko

25 September 2021

marginal

$X = \#$ of 'H' after 2 tosses

$E[X] = 2\cdot\frac{1}{4} + 1\cdot\frac{1}{4} + 1\cdot\frac{1}{4} + 0\cdot\frac{1}{4} = \frac{4}{4} = 1$




$$\begin{cases} E[X \mid \mathcal{F}_1] = 2\cdot\frac{1}{2} + 1\cdot\frac{1}{2} = 3/2 \\ E[X \mid \mathcal{F}_2] = 1\cdot\frac{1}{2} + 0\cdot\frac{1}{2} = 1/2 \end{cases}$$

Conditional expectation

$$E[X \mid \mathcal{F}_i] = \begin{cases} 3/2 & \text{if } i=1 \quad 1/2 \\ 1/2 & \text{if } i=2 \quad 1/2 \end{cases}$$

Tower's Property: $E\left[E[X \mid \mathcal{F}_i]\right] = \frac{3}{2}\cdot\frac{1}{2} + \frac{1}{2}\cdot\frac{1}{2} = \frac{4}{4} = 1$

$E[E[X \mid \mathcal{F}_i]] = E[X]$ !

$E[X]$

$\sum_{i=1}^{+\infty} x_i \, f_{X|Y}$

Filtration

## Lecture

$f_{X|Y} \rightarrow E[X|Y] = \int_{-\infty}^{+\infty} x \cdot f_{X|y} \, dx \Rightarrow E[X|Y] = g(y)$

$E[X|Y]$

- It all starts with the definition of conditional probability: $P(A|B) = P(AB)/P(B)$.
- If $X$ and $Y$ are jointly discrete random variables, we can use this to define a probability mass function for $X$ *given* $Y = y$.
- That is, we write $p_{X|Y}(x|y) = P\{X = x|Y = y\} = \frac{p(x,y)}{p_Y(y)}$.
- In words: first restrict sample space to pairs $(x, y)$ with given $y$ value. Then divide the original mass function by $p_Y(y)$ to obtain a probability mass function on the restricted space.
- We do something similar when $X$ and $Y$ are continuous random variables. In that case we write $f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)}$.
- Often useful to think of sampling $(X, Y)$ as a two-stage process. First sample $Y$ from its marginal distribution, obtain $Y = y$ for some particular $y$. Then sample $X$ from its probability distribution *given* $Y = y$.
- Marginal law of $X$ is weighted average of conditional laws.

# Lecture

$\boxed{Z = X + 5}$

- Let $X$ be value on one die roll, $Y$ value on second die roll, and write $Z = X + Y$.
- What is the probability distribution for $X$ given that $Y = 5$?
- Answer: uniform on $\{1, 2, 3, 4, 5, 6\}$. ✓
- What is the probability distribution for $Z$ given that $Y = 5$?
- Answer: uniform on $\{6, 7, 8, 9, 10, 11\}$.
- What is the probability distribution for $Y$ given that $Z = 5$?
- Answer: uniform on $\{1, 2, 3, 4\}$.

Since $X, Y$ - ind-t $\Rightarrow P(X|Y) = \frac{P(X \cap Y)}{P(Y)} = \frac{P(X) \cdot P(Y)}{P(Y)} = P(X)$

| X | Y |
|---|---|
| 1 | 4 |
| 2 | 3 |
| 3 | 2 |
| 4 | 1 |
| 5 | |
| 5 | |

## Lecture

- Now, what do we mean by $E[X|Y = y]$? This should just be the expectation of $X$ in the conditional probability measure for $X$ given that $Y = y$.
- Can write this as $E[X|Y = y] = \sum_x xP\{X = x|Y = y\} = \sum_x xp_{X|Y}(x|y)$.
- Can make sense of this in the continuum setting as well.
- In continuum setting we had $f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)}$. So $E[X|Y = y] = \int_{-\infty}^{\infty} x \frac{f(x,y)}{f_Y(y)} dx$

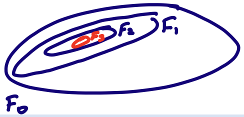
## Lecture

$X \sim U[1,2,3,4,5,6]$

- Let $X$ be value on one die roll, $Y$ value on second die roll, and write $Z = X + Y$.
- What is $E[X|Y=5]$? $= 1\cdot\frac{1}{6} + 2\cdot\frac{1}{6} + 3\cdot\frac{1}{6} + 4\cdot\frac{1}{6} + 5\cdot\frac{1}{6} + \frac{1}{6}\cdot 6 = \ldots$
- What is $E[Z|Y=5]$? $= 6\cdot\frac{1}{6} + 7\cdot\frac{1}{6} + \ldots + 11\cdot\frac{1}{6} = E[X+5 \mid Y=5] = \ldots$
- What is $E[Y|Z=5]$? $= 1\cdot\frac{1}{4} + 2\cdot\frac{1}{4} + 3\cdot\frac{1}{4} + 4\cdot\frac{1}{4} = \ldots$

$F_0$ - not precise

$F_3 < F_0$ more precise




## Lecture

$E\Big[E\big[E[E(X|\mathcal{F}_3)|\mathcal{F}_2]\,|\,\mathcal{F}_1\big]\,\Big|\,\mathcal{F}_0\Big] = E[X]$ ! $\mathcal{F}_0 \supseteq \mathcal{F}_1 \supseteq \mathcal{F}_2 \supseteq \mathcal{F}_3$

- Can think of $E[X|Y]$ as a function of the random variable $Y$. When $Y = y$ it takes the value $E[X|Y = y]$.
- So $E[X|Y]$ is itself a random variable. It happens to depend only on the value of $Y$.
- Thinking of $E[X|Y]$ as a random variable, we can ask what *its* expectation is. What is $E[E[X|Y]]$?
- ! **Very useful fact:** $E[E[X|Y]] = E[X]$. Tower's Prop-ty
- In words: what you expect to expect $X$ to be *after learning* $Y$ is same as what you *now* expect $X$ to be.
- Proof in discrete case: $E[X|Y = y] = \sum_x xP\{X = x|Y = y\} = \sum_x x\frac{p(x,y)}{p_Y(y)}$.
- Recall that, in general, $E[g(Y)] = \sum_y p_Y(y)g(y)$.
- $E[E[X|Y = y]] = \sum_y p_Y(y)\sum_x x\frac{p(x,y)}{p_Y(y)} = \sum_x\sum_y p(x,y)x = E[X]$.

- Definition:
  $\mathrm{Var}(X|Y) = E\big[(X - E[X|Y])^2|Y\big] = E\big[X^2 - E[X|Y]^2|Y\big]$.
- $\mathrm{Var}(X|Y)$ is a random variable that depends on $Y$. It is the variance of $X$ in the conditional distribution for $X$ given $Y$.
- Note $E[\mathrm{Var}(X|Y)] = E[E[X^2|Y]] - E[E[X|Y]^2|Y] = E[X^2] - E[E[X|Y]^2]$.
- If we subtract $E[X]^2$ from first term and add equivalent value $E[E[X|Y]]^2$ to the second, RHS becomes $\mathrm{Var}[X] - \mathrm{Var}[E[X|Y]]$, which implies following:
- **Useful fact:** $\mathrm{Var}(X) = \mathrm{Var}(E[X|Y]) + E[\mathrm{Var}(X|Y)]$.
- One can discover $X$ in two stages: first sample $Y$ from marginal and compute $E[X|Y]$, then sample $X$ from distribution given $Y$ value.
- Above fact breaks variance into two parts, corresponding to these two stages.

## Lecture

$E[Z|X] = E[X+Y|X] = \underbrace{E[X|X]}_{=\boxed{X}} + E[Y|X] = X + \underbrace{E[Y|X]}_{= X + E[Y]}$

- Let $X$ be a random variable of variance $\sigma_X^2$ and $Y$ an independent random variable of variance $\sigma_Y^2$ and write $Z = X + Y$. Assume $E[X] = E[Y] = 0$.
- What are the covariances $\underbrace{\mathrm{Cov}(X, Y)}_{=0}$ and $\mathrm{Cov}(X, Z)$?

  $\mathrm{Cov}(X, X+Y) = \mathrm{Var}\, X + \underbrace{\mathrm{Cov}(X,Y)}_{=0}$
- How about the correlation coefficients $\rho(X, Y)$ and $\rho(X, Z)$?
- What is $E[Z|X]$? And how about $\mathrm{Var}(Z|X)$?
- Both of these values are functions of $X$. Former is just $X$. Latter happens to be a constant-valued function of $X$, i.e., happens not to actually depend on $X$. We have $\mathrm{Var}(Z|X) = \sigma_Y^2$.
- Can we check the formula $\mathrm{Var}(Z) = \mathrm{Var}(E[Z|X]) + E[\mathrm{Var}(Z|X)]$ in this case?

## Lecture

- Sometimes think of the expectation $E[Y]$ as a "best guess" or "best predictor" of the value of $Y$.
- It is best in the sense that at among all constants $m$, the expectation $E[(Y-m)^2]$ is minimized when $m = E[Y]$.
- But what if we allow non-constant predictors? What if the predictor is allowed to depend on the value of a random variable $X$ that we can observe directly?
- Let $g(x)$ be such a function. Then $E[(y-g(X))^2]$ is minimized when $g(X) = E[Y|X]$.

$$\boxed{\begin{gathered}\min \mathrm{Var}(X)\\ E[X] = m\end{gathered}}$$

# Lecture

- Toss 100 coins. What's the conditional expectation of the number of heads given the number of heads among the first fifty tosses?
- What's the conditional expectation of the number of aces in a five-card poker hand given that the first two cards in the hand are aces?

# Lecture

*Conditional expectation, $\mathbb{E}(X \mid Y)$, is a random variable with randomness inherited from $Y$, not $X$.*

***Example:*** Suppose $Y = \begin{cases} 1 & \text{with probability } 1/8\,, \\ 2 & \text{with probability } 7/8\,, \end{cases}$

and $\quad X \,|\, Y = \begin{cases} 2Y & \text{with probability } 3/4\,, \\ 3Y & \text{with probability } 1/4\,. \end{cases}$

$$E[X|Y] = 2Y \cdot \frac{3}{4} + 3Y \cdot \frac{1}{4} = \frac{7Y}{4} = \frac{7}{4} \cdot Y$$

$$E[\,E[X|Y]\,] = E\left[\frac{7}{4} \cdot Y\right] = \frac{7}{4} \cdot \left(1 \cdot \frac{1}{8} + 2 \cdot \frac{7}{8}\right) = \frac{7}{4} \cdot \frac{15}{8}.$$

## Conditional variance

The conditional variance is similar to the conditional expectation.

- $\text{Var}(X \mid Y = y)$ is the variance of $X$, when $Y$ is fixed at the value $Y = y$.
- $\text{Var}(X \mid Y)$ is a random variable, giving the variance of $X$ when $Y$ is fixed at a value to be selected randomly.

*Definition:* Let $X$ and $Y$ be random variables. The **conditional variance of $X$, given $Y$,** is given by

$$\boxed{\text{Var}(X \mid Y) = \mathbb{E}(X^2 \mid Y) - \left\{\mathbb{E}(X \mid Y)\right\}^2 = \mathbb{E}\left\{(X - \mu_{X \mid Y})^2 \mid Y\right\}}$$

# Lecture

If all the expectations below are finite, then for ANY random variables $X$ and $Y$, we have:

i) $\boxed{\mathbb{E}(X) = \mathbb{E}_Y\Big(\mathbb{E}(X \mid Y)\Big)}$ *Law of Total Expectation.*

*Note that we can pick any r.v. $Y$, to make the expectation as easy as we can.*

ii) $\mathbb{E}(g(X)) = \mathbb{E}_Y\Big(\mathbb{E}(g(X) \mid Y)\Big)$ *for any function $g$.*

iii) $\boxed{\mathrm{Var}(X) = \mathbb{E}_Y\Big(\mathrm{Var}(X \mid Y)\Big) + \mathrm{Var}_Y\Big(\mathbb{E}(X \mid Y)\Big)}$

*Law of Total Variance.*

# Lecture

## 1. Swimming with dolphins

Fraser runs a dolphin-watch business. Every day, he is unable to run the trip due to bad weather with probability $p$, independently of all other days. Fraser works every day except the bad-weather days, which he takes as holiday.

Let $Y$ be the number of consecutive days Fraser has to work between bad-weather days. Let $X$ be the total number of customers who go on Fraser's trip in this period of $Y$ days. Conditional on $Y$, the distribution of $X$ is

$$(X \mid Y) \sim \text{Poisson}(\mu Y).$$

(a) Name the distribution of $Y$, and state $\mathbb{E}(Y)$ and $\text{Var}(Y)$.

(b) Find the expectation and the variance of the number of customers Fraser sees between bad-weather days, $\mathbb{E}(X)$ and $\text{Var}(X)$.

(a) *Let 'success' be 'bad-weather day' and 'failure' be 'work-day'.*
*Then* $\mathbb{P}(\textit{success}) = \mathbb{P}(\textit{bad-weather}) = p$.
$Y$ *is the number of failures before the first success.*
*So*

$$Y \sim \text{Geometric}(p).$$

*Thus*

$$\begin{aligned} \mathbb{E}(Y) &= \frac{1-p}{p}, \\ \text{Var}(Y) &= \frac{1-p}{p^2}. \end{aligned}$$

# Lecture

(b) *We know* $(X \mid Y) \sim$ *Poisson*$(\mu Y)$*: so*

$$\mathbb{E}(X \mid Y) = \text{Var}(X \mid Y) = \mu Y.$$

*By the Law of Total Expectation:*

$$\begin{aligned} \mathbb{E}(X) &= \mathbb{E}_Y\Big\{\mathbb{E}(X \mid Y)\Big\} \\ &= \mathbb{E}_Y(\mu Y) \\ &= \mu\mathbb{E}_Y(Y) \end{aligned}$$

$$\therefore \quad \mathbb{E}(X) = \frac{\mu(1-p)}{p}.$$

*By the Law of Total Variance:*

$$\begin{aligned} \text{Var}(X) &= \mathbb{E}_Y\Big(\text{Var}(X \mid Y)\Big) + \text{Var}_Y\Big(\mathbb{E}(X \mid Y)\Big) \\ &= \mathbb{E}_Y\Big(\mu Y\Big) + \text{Var}_Y\Big(\mu Y\Big) \\ &= \mu\mathbb{E}_Y(Y) + \mu^2\,\text{Var}_Y(Y) \\ &= \mu\left(\frac{1-p}{p}\right) + \mu^2\left(\frac{1-p}{p^2}\right) \\ &= \frac{\mu(1-p)(p+\mu)}{p^2}. \end{aligned}$$

## 2. Randomly stopped sum


This model arises very commonly in stochastic processes. A random number $N$ of events occur, and each event $i$ has associated with it some cost, penalty, or reward $X_i$. The question is to find the mean and variance of the total cost / reward: $T_N = X_1 + X_2 + \ldots + X_N$.
The difficulty is that the number $N$ of terms in the sum is itself random.

$T_N$ is called a ***randomly stopped sum: it is a sum of $X_i$'s, randomly stopped at the random number of $N$ terms.***

***Example:*** Think of a cash machine, which has to be loaded with enough money to cover the day's business. The number of customers per day is a random number $N$. Customer $i$ withdraws a random amount $X_i$. The total amount withdrawn during the day is a randomly stopped sum: $T_N = X_1 + \ldots + X_N$.

$P(N \leq \text{time} \mid T_N \geq 1000)$

# Lecture

## Cash machine example

The citizens of Remuera withdraw money from a cash machine according to the following probability function ($X$):

| Amount, $x$ (\$) | 50 | 100 | 200 |
|---|---|---|---|
| $\mathbb{P}(X = x)$ | 0.3 | 0.5 | 0.2 |

The number of customers per day has the distribution $N \sim \text{Poisson}(\lambda)$.

Let $T_N = X_1 + X_2 + \ldots + X_N$ be the total amount of money withdrawn in a day, where each $X_i$ has the probability function above, and $X_1, X_2, \ldots$ are independent of each other and of $N$.
$T_N$ is a randomly stopped sum, stopped by the random number of $N$ customers.

(a) Show that $\mathbb{E}(X) = 105$, and $\text{Var}(X) = 2725$.

(b) Find $\mathbb{E}(T_N)$ and $\text{Var}(T_N)$: the mean and variance of the amount of money withdrawn each day.

# Lecture

*Similarly,*

$$
\begin{aligned}
\mathrm{Var}(T_N \,|\, N) &= \mathrm{Var}(X_1 + X_2 + \ldots + X_N \,|\, N) \\
&= \mathrm{Var}(X_1 + X_2 + \ldots + X_N) \\
&\qquad\qquad \textit{where } N \textit{ is now considered constant;} \\
&\qquad\qquad \textit{(because all } X_i\textit{'s are independent of } N\textit{)} \\
&= \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \ldots + \mathrm{Var}(X_N) \\
&\qquad \textit{(we DO need independence of } X_i\textit{'s for this)} \\
&= N \times \mathrm{Var}(X) \quad \textit{(because all } X_i\textit{'s have same variance, } \mathrm{Var}(X)\textit{)} \\
&= 2725N.
\end{aligned}
$$

# Lecture

*So*

$$\begin{aligned}
\mathbb{E}(T_N) &= \mathbb{E}_N \left\{ \mathbb{E}(T_N \mid N) \right\} \\
&= \mathbb{E}_N(105N) \\
&= 105\mathbb{E}_N(N) \\
&= 105\lambda,
\end{aligned}$$

*because* $N \sim$ *Poisson*$(\lambda)$ *so* $\mathbb{E}(N) = \lambda$.

*Similarly,*

$$\begin{aligned}
\textit{Var}(T_N) &= \mathbb{E}_N \left\{ \textit{Var}(T_N \mid N) \right\} + \textit{Var}_N \left\{ \mathbb{E}(T_N \mid N) \right\} \\
&= \mathbb{E}_N \{2725N\} + \textit{Var}_N \{105N\} \\
&= 2725\mathbb{E}_N(N) + 105^2 \textit{Var}_N(N) \\
&= 2725\lambda + 11025\lambda \\
&= 13750\lambda,
\end{aligned}$$

*because* $N \sim$ *Poisson*$(\lambda)$ *so* $\mathbb{E}(N) = \textit{Var}(N) = \lambda$.

**General result for randomly stopped sums:**

Suppose $X_1, X_2, \ldots$ each have the same mean $\mu$ and variance $\sigma^2$, and $X_1, X_2, \ldots,$ and $N$ are mutually independent. Let $T_N = X_1 + \ldots + X_N$ be the randomly stopped sum. By following similar working to that above:

$$\boxed{\begin{aligned} \mathbb{E}(T_N) = \mathbb{E}\left\{\sum_{i=1}^{N} X_i\right\} &= \mu\,\mathbb{E}(N) \\ \mathrm{Var}(T_N) = \mathrm{Var}\left\{\sum_{i=1}^{N} X_i\right\} &= \sigma^2\,\mathbb{E}(N) + \mu^2\,\mathrm{Var}(N). \end{aligned}}$$

Thank you for your attention!

See next week!