STAT 333 Lecture Blog Winter 2012 (MWF)

Lecture Blog

Back to Forums & Topics List

Compose Refresh Mark All Read Hide Search Unsubscribe to Topic Subscriptions Settings View:

I plan to use this space to post brief summaries of what we talked about in lectures. You can use this space to clarify concepts, respond to take-home questions if I ask any, and see what you missed if you weren't in class for some reason.

Lecture 36 - April 2

Diana Katherine Skrzydlo - Apr 2, 2012 10:53 AM
Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

Today was a review class. Thanks for a great term everyone!

Lecture 35 - March 30

Diana Katherine Skrzydlo - Mar 30, 2012 12:43 PM
Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

Today we looked at some more complex examples of birth-and-death processes.

We can modify the single-server queue in several ways, including adding a maximum system capacity or adding more servers.

We looked at an example of multiple machines working (and being fixed) at the same time, which is similar but not identical to the assignment question 5.

Finally we looked at a linear growth model (with both birth and death rates being linear functions of the population size)

Lecture 34 - March 28

0 unread of 1 messages - 1

Diana Katherine Skrzydlo - Mar 30, 2012 12:41 PM
Mark Unread
[Reply]
More actions...

author(s)

Today we derived the mathematical model for a birth-and-death process. We can find the total rate out of each state using the alarm clock lemma. $v_0 = \lambda_0$, and otherwise $v_i = \lambda_i + \mu_i$. Because there are only two possible places we can go from state i (i+1 or i-1), most of the instantaneous transition probabilities are 0, and hence so is most of the **Q** matrix. It turns out that **Q** is so simple that we can explicitly solve for the equilibrium probabilities. $\pi_0 = 1/(1 + \sum (\lambda_0 \ldots \lambda_{j-1}/\mu_1 \ldots \mu_j))$, if the sum converges, and all the other π_j's are a multiple of π_0.

We looked at a couple of examples - the poisson process (pure birth), pure death process, and a single-server queue.

Lecture 33 - March 26

Diana Katherine Skrzydlo - Mar 26, 2012 10:53 AM
Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

Today we looked at a couple of examples of continuous MCs.

We started with a 2-state chain, and were able to write down the general form of **Q** and use it to find **P** (very simple), $\boldsymbol{\pi}$, and through the Kolmogorov equations, **P(t)**. A handout is posted with some of the technical details. We then checked that **P(t)** gave us the same results for $\boldsymbol{\pi}$ (both rows as t -> ∞) and **Q** (the derivative at t=0).

We also did a 3-state chain example where we started with **Q**, and looked at why certain states had higher equilibrium probabilities than others.

Finally, we defined what it means for a cts MC to be a birth and death process, and talked briefly about situations where that model is appropriate. We'll see more details and examples on Wednesday.

Someone asked a question after class - is it possible to have an absorbing state in a cts MC? If so, what would the values in the various matrices be for it?

A: If state i is absorbing, the process cannot leave it, so $v_i = 0$ and $P_{ij} = 0$ for $j \neq i$. The entire i^{th} row of the **Q** matrix would be 0, and for **P(t)**, $P_{ii}(t) = 1$, $P_{ij}(t) = 0$ for all t. The only weird thing would be the instantaneous transition probability P_{ii}. Normally we say that is 0, but if i is absorbing it has value 0/0 which is indeterminate. To keep **P** being a stochastic matrix, a useful convention would be to have $P_{ii} = 1$ if i is an absorbing state. Great question!

Lecture 32 - March 23

Diana Katherine Skrzydlo - Mar 26, 2012 10:45 AM

Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

Today was all about the matrix **Q**. We showed that it has its off-diagonal elements $q_{ij} = v_i*P_{ij}$ and its diagonal elements $q_{ii} = -v_i$. We can use **Q** to solve for the equilibrium distribution $\boldsymbol{\pi}$, to recover the v_i's and P_{ij}'s (if we didn't already have them), and to actually obtain the matrix family **P(t)** by using the Kolmogorov Forward/Backward Equations, which we proved for the backward case. Note: You will not have to solve matrix differential equations, but you should know how to set them up (given a **Q** matrix, show the pointwise equations the $P_{ij}(t)$'s must solve.)

See here for an awesome video from 2010's YouTube assignment explaining all about **Q**.

Lecture 31 - March 21

Diana Katherine Skrzydlo - Mar 26, 2012 10:42 AM

Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

Today we started by talking about the different types of movement in a cts MC. We showed that the waiting times until the process leave state i must be memoryless (and hence exponential with some rate v_i), and defined the embedded chain (or instantaneous transition) probabilities P_{ij}, which then gave us the instantaneous transition rate from i to j, $q_{ij} = v_iP_{ij}$.

We then defined the matrix **Q** as the derivative at 0 (from above) of **P(t)**. Q has all its off-diagonal elements non-negative, but the rows sum to 0, not 1 (so the diagonal elements are negative.) It turns out that **Q** contains all the information about the MC, including the equilibrium distritbution, but we don't usually get it from P(t), since that is hard to find.

Lecture 30 - March 19

Diana Katherine Skrzydlo - Mar 19, 2012 12:34 PM

Mark Unread

0 unread of 1 messages - 1 author(s)

[Reply]
More actions...

Today we started chapter 6 - continuous time Markov Chains.

We started by speaking in general terms about the movement in a cts MC, which is governed by both the time spent in each state (exponential) and the probabilities of where the process moves when it leaves a state (instantaneous tpm **P**). We defined the continuous version of the Markov property, and the infinite family of transition matrices **P(t)**. The C-K equations hold just like they do in the discrete case, but unfortunately we can't truly capture the continuous movements with a single matrix. Looking at the long run behaviour, everything is the same as the discrete case (we keep all our definitions except nothing is periodic anymore) and the vector $\boldsymbol{\pi}$ exists uniquely if the chain is irreducible, but we don't have an easy way to solve for $\boldsymbol{\pi}$.

Next time we'll see a matrix **Q** that solves all these problems!

Lecture 29 - March 16

Diana Katherine Skrzydlo - Mar 19, 2012 12:31 PM
Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

Today we finished off chapter 5 by talking about some of the properties of the poisson process.

We talked about thinning, which is where we can create two (or more!) new poisson processes from a single one by classifying each event as one of two (or more) types. The resulting process of only type 1 events is still a poisson process. For the proof, we can look at the result from tutorial 3 or test 1.

We also looked at the conditional distribution of the event times, given we know how many there are in a time interval. They turn out to be each uniform(0,t) and independent. We proved the case for a single event, and someone afterwards suggested we could prove the general case using strong induction rather than order statistics. This might be possible, I'll think about it.

Finally, the number of events in a subinterval of length s (again conditional on n events in (0, t)) is binomial (n, s/t). We proved this mathematically and justified it logically.

Lecture 28 - March 14 (Pi day!)

Diana Katherine Skrzydlo - Mar 14, 2012 5:08 PM
Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

Today we talked all about the Poisson process.

We first defined a counting process, and talked about stationary and independent increments, then described 4 different (but equivalent) definitions of the Poisson process. We can easily prove some of the equivalences, in fact some you have already seen. We also looked at a sample path (or trajectory) of a Poisson process, which is an increasing right-continuous step function. We'll talk about a few more properties of the Poisson process on Friday.

Lecture 27 - March 12

Diana Katherine Skrzydlo - Mar 12, 2012 9:35 AM
Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

Today we started Chapter 5 by talking about the exponential distribution.

This is the only continuous distribution that has the memoryless property, and hence must be the distribution for how long we spend in each state in a continuous Markov Chain. We proved the Alarm Clock Lemma which says that, for a collection of n independent $T_i \sim \exp(\lambda_i)$,

1. $T = \min\{T_i\} \sim \exp(\Sigma\lambda_i)$
2. $P(T_j = T) = \lambda_j / \Sigma\lambda_i$
3. The event $T_j = T$ is independent of the value of T

Lecture 26 - March 9

Diana Katherine Skrzydlo - Mar 9, 2012 12:00 PM
Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

Today we looked at a pretty involved example where we had a markov chain with multiple closed classes.

From the transition probability matrix **P** we were able to determine the classes (using a state diagram) and rewrite P in simplified form by reordering the states. Within each class we found the mini-equilibrium distribution (we did this using logic but you could also solve $\underline{\pi}_C \mathbf{P}_C = \underline{\pi}_C$ subject to the elements of $\underline{\pi}_C$ adding to 1). Then for this chain, there are actually an infinite number of stationary distributions, each a linear combination of the two equilibrium distributions for the closed classes. We can write the general form as $\alpha\underline{\pi}_{C1} + (1-\alpha)\underline{\pi}_{C2}$ for any real α between 0 and 1.

We then talked about absorption probabilities (the final topic in discrete MCs) from transient states into closed classes. In general, we can find $A_{k,Cj}$ as the probability of going directly from k into Cj plus the probability of going somewhere else (transient), then from there being absorbed into Cj. We found the probabilities for our example. The actual equilibrium reached in these types of markov chains depends on the starting state. I will post some R code and results so you can see equilibrium behaviour for various MCs.

Lecture 25 - March 7

Diana Katherine Skrzydlo - Mar 9, 2012 11:50 AM
Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

Today we looked in more detail at finding equilibrium for a markov chain.

Any vector $\underline{\pi}$ that satisfies $\underline{\pi}\,\mathbf{P} = \underline{\pi}$ and has non-negative elements that sum to 1 is called a stationary distribution of the markov chain. It turns out that $\underline{\pi}$ is unique and guaranteed to exist if the MC is irreducible and the states are positive recurrent. If the MC is also aperiodic, then as $n \to \infty$ the matrix $\mathbf{P}^{(n)}$ will approach a matrix $\prod$ which has all its rows equal to $\underline{\pi}$. This means that the state the MC is in in the long run does not depend on the starting state. We can interpret π_i as the long-run proportion of time the process spends in state i, and $1/\pi_i$ is the average number of steps to return to state i.

If a state is transient or null recurrent, the equilibrium probability for that state is 0.

If a MC has two or more closed classes, then unfortunately $\underline{\pi}$ is not unique, but we can find a mini-equilibrium distribution within each closed class.

Lecture 24 - March 5

Diana Katherine Skrzydlo - Mar 9, 2012 11:44 AM
Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

We again talked about classes, and proved several results:
Open classes contain transient states
Finite closed classes contain positive recurrent states
Infinite closed classes contain all of any type of state

We stated but did not prove that the period and type of state is the same for all states in a class. We examined some simple examples of finding the period of a class based on the possible one-step transitions.

We also defined the row vector $\underline{p}_n$ which contains the probabilities of the chain being in each state at step n. Because of the Chapman-Kolmogorov equations, we can find that $\underline{p}_{n+1} = \underline{p}_n \mathbf{P}$. In some cases, the markov chain will approach an equilibrium in the long run, and in that case we will have $\underline{p}_n = \underline{p}_{n+1} = \underline{\pi}$, so $\underline{\pi}\,\mathbf{P} = \underline{\pi}$.

Lecture 23 - March 2

Diana Katherine Skrzydlo - Mar 9, 2012 11:38 AM
Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

Today we talked about how renewal theory applies to Markov Chains.

We steal all the terminology for events in renewal theory and apply it to the states in a MC. There is a handout on this material posted. We looked at some examples and found that some chains have all the states with the same behaviour but some have states with different behaviour in the long run. This is because of...

Classes in markov chains. (There is a handout on this material posted too.) We defined what it means for states to be accessible from one another and to communicate - there is a path of any length from one state to the other, and a (possibly different) past back. Communication splits the state space into classes where two states are in the same class if and only if they communicate. There are two types of classes, open (which you can leave but then not come back) and closed (which you cannot leave).

Lecture 22 - Feb 29

Diana Katherine Skrzydlo - Mar 9, 2012 11:33 AM
Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

We looked at some simple examples of Markov chains, including a two-state chain (commonly used for the weather), a chain where the win or loss of a team depends on the last 2 outcomes (we need 4 states rather than 2 for this to satisfy the Markov property), and the various modifications of the random walk that we've seen.

We defined the n-step transition probabilities $P^{(n)}_{ij}$ and derived the Chapman-Kolmogorov equations which allow us to use the matrix **P** to obtain any transition probability for any number of steps in the future.

We started to look at renewal theory applied to MCs.

Lecture 21 - Feb 27

Diana Katherine Skrzydlo - Mar 9, 2012 11:25 AM

0 unread of 1 messages - 1 author(s)

Mark Unread
[Reply]
More actions...

Today we started talking about Markov Chains (chapter 4 in the textbook).

We started by looking at the random walk on a circle, which is difficult to analyze using renewal theory techniques. We defined the conditions for a markov chain - discrete stochastic process with a discrete state space and the markov property, which is that the future can only depend on the present, not the past.

We defined the one-step transition probabilities P_{ij} and the matrix of these values **P**, which is known as a stochastic matrix and has all elements non-negative and all rows sum to 1.

We brainstormed some situations where markov chains can be used as models.

Lecture 20 - Feb 17

Diana Katherine Skrzydlo - Feb 17, 2012 11:21 AM

0 unread of 1 messages - 1 author(s)

Mark Unread
[Reply]
More actions...

Today we finished up renewal theory by talking about the event λ_{01} = "get to 1 starting from 0" which is delayed renewal. It turns out to be guaranteed to occur if $p \geq 1/2$ but not if $p < 1/2$. Furthermore, its average waiting time is finite if $p > 1/2$.

We can extend these results to events for getting from any state to any state, including going in the negative direction by reversing the roles of p and q.

Lastly, we talked briefly about gambler's ruin, a modification of the random walk with barriers at 0 and k. More details of the derivation of the solution are posted in a handout, and an example.

Have a great reading week!

Lecture 19 - Feb 16 (in tutorial)

Diana Katherine Skrzydlo - Feb 16, 2012 11:38 PM
Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

Today we talked more about the random walk. In particular, we analyzed the event λ_{00} = "return to 0" in a number of ways. We found the renewal sequence (in this case it was periodic with period 2 so all the odd terms in the sequence are 0) but it was hard to evaluate the limit and use the renewal theorem. Instead we summed the renewal sequence to get $E[V_{00}]$ by approximating it and showing it diverges if p=1/2 and converges if $p \neq 1/2$. Hence the event "return to 0" is transient if the walk is biased and recurrent if the walk is balanced.

We found more detailed information by actually deriving the pgf of T_{00} (the waiting time until the first occurrence of λ_{00}. We used some technical lemmas and got $F_{00}(s) = 1 - \sqrt{(1 - 4pqs^2)}$. This showed us that is null recurrent if p = 1/2, and also gave us a way to find the exact value for $E[V_{00}]$.

Next time we'll see other events in this process as well as a common modification, the gambler's ruin model.

Lecture 18 - Feb 15

Diana Katherine Skrzydlo - Feb 16, 2012 11:32 PM
Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

Today we explained the difference between the waiting time for the delayed renewal event "22" and the renewal event "24". Then we talked about the renewal theorem (both periodic and aperiodic) and how that can be used to easily find the average waiting time for a pattern from the renewal sequence.

We introduced the random walk, a discrete stochastic process where each time the process moves up by 1 or down by 1 independently of other steps. A handout on the setup and basic results is posted.

Lecture 17 - Feb 13

Diana Katherine Skrzydlo - Feb 13, 2012 10:18 AM

Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

Today we used the renewal relation to find out all the information about a renewal event "24" in rolls of a fair die. We were able to find the renewal sequence and get the pgf, from which we found the event was recurrent, had mean 36 (which makes sense), and get the probabilities of it occurring for the first time on trials 2, 3, and 4 (which we could also come up with a logical explanation for).

We then talked about what happens in the delayed renewal case. We get the delayed renewal sequence d_0, d_1, d_2 etc which is the same as r_n except $d_0=0$ and the generating function of that is $D_\lambda(s)$. We also have an associated renewal event $\not\lambda$ which is conditional on λ already having occurred. We get the associated renewal sequence $\not r_0$, $\not r_1$, $\not r_2$ etc and $R_{\not\lambda}(s)$. Then the pgf of the first waiting time for the delayed renewal event is $F_\lambda(s) = D_\lambda(s)/R_{\not\lambda}(s)$. We did an example with "22" in rolls of a die, and surprisingly found that $E[T_\lambda] = 42$.

Lecture 16 - Feb 10

Diana Katherine Skrzydlo - Feb 12, 2012 10:34 PM

Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

Today we talked more about the random variable V_λ, which can tell us about whether the event λ is transient or recurrent, but unfortunately not much else. We also introduced the renewal sequence, an easy to obtain collection of probabilities with $r_0 = 1$, $r_n = P(\lambda$ occurs on trial n). We can use the renewal sequence in two ways - to calculate $E[V_\lambda]$ by summing rn from n=1 to ∞, or by finding the generating function of the r_n's, $R_\lambda(s)$ and using the renewal relation to find the pgf of T_λ, $F_\lambda(s)$. We'll see some examples next time, and how this changes with a delayed renewal event.

Lecture 15 - Feb 8

Diana Katherine Skrzydlo - Feb 8, 2012 5:18 PM

Mark Unread

[Reply]

More actions...

0 unread of 1 messages - 1 author(s)

Today we classified renewal events into transient or recurrent based on whether or not they are guaranteed to occur. We also further split up recurrent events (ones which are guaranteed to occur) into positive recurrent (average waiting time is finite) and null recurrent (average waiting time is infinite). These classifications correspond exactly to the definitions of improper, proper, short proper, and null proper random variables.

We then defined a new random variable V_λ which represents the number of times λ occurs. We derived the pmf and it turns out that $V_\lambda + 1 \sim Geo(1-f_\lambda)$. (I wrote that incorrectly on the board as $V_\lambda \sim 1 + Geo(1-f_\lambda)$, please correct your notes, sorry!)

Lecture 14 - Feb 6

Diana Katherine Skrzydlo - Feb 8, 2012 5:14 PM

Mark Unread

[Reply]

More actions...

0 unread of 1 messages - 1 author(s)

Today we started talking about renewal theory. We defined stochastic process (and looked at a few examples), an event λ, the waiting time T_λ, the probability that T_λ is finite f_λ, and the inter-event waiting times $T_\lambda^{(k, k+1)}$. We defined renewal events (all T_λ's iid) and delayed renewal events (the first one is different), and looked at a brief example of both kinds. All these definitions are summarized in a handout.

Lecture 13 - Feb 1

Diana Katherine Skrzydlo - Feb 1, 2012 9:55 AM

Mark Unread

[Reply]

More actions...

0 unread of 1 messages - 1 author(s)

Today we finished up talking about probability generating functions. The first test covers the material up to the end of today. Please don't forget there is NO CLASS THIS FRIDAY.

We talked about what we can do with pgfs:
- determine if a rv is proper or improper
- derive the pmf of the rv (can be tedious if you use Taylor series, but can always be done)
- find the mean if the rv is proper by taking the first derivative at s=1
- find the variance
- prove stuff easily

We also looked at the multiplicative property which holds when the random variables are independent, and how that can help us with compound random variables.

Please read the handouts and let me know if any of the examples are unclear. See you next week!

Lecture 12 - Jan 30

Diana Katherine Skrzydlo - Jan 30, 2012 9:34 AM
Mark Unread
[Reply]
More actions...

0 unread of 3 messages - 2 author(s)

Today we started talking about generating functions, and in particular probability generating functions. This material is not in the textbook so make sure you look at the handouts posted. We defined a generating function (when it exists), looked at how to find the coefficients using a Taylor series expansion or more useful techniques such as combining other generating functions using addition/subtraction/integration/differentiation/multiplication. We defined a probability generating function $G_X(s)$and showed that it will always converge absolutely on the interval (-1, 1). We can check whether X is proper by checking if GX(1) = 1. If it's <1, X is improper.

Collapse Replies

Re: Lecture 12 - Jan 30

Student - Jan 30, 2012 7:55 PM
Mark Unread
[Reply]
More actions...

0 unread of 2 messages - 2 author(s)

Will the generating function still converge on at least (-1,1) if the RV is improper?

Collapse Replies

Re: Lecture 12 - Jan 30

[Diana Katherine Skrzydlo] - Jan 30, 2012 10:45 PM
Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

Yes. In fact, at s=1 it will converge to a value strictly less than 1 if X is improper. My convention is, when I write the sum as n goes from 0 to infinity, not to include the actual term at infinity. But even if you do, if |s| <= 1, you would get s^infinity equalling 0, so the p_infinity term will not contribute.

Lecture 11 - Jan 27

[Diana Katherine Skrzydlo] - Jan 27, 2012 9:59 AM
Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

Today we finished Chapter 3 by talking about compound random variables, which are a sum of a random number of independent and identically distributed random variables. $S = X_1 + X_2 + ... + X_N$

We derived the mean and variance of S by conditioning on N and using double averaging and the conditional variance formula. We get $E[S] = E[X]E[N]$ and $Var(S) = Var(X)E[N] + E[X]^2Var(N)$. We saw a quick example.

Questions afterward:

Can we condition on X instead of N to derive these formulae?

No, because all the X_i's are not the same. If we condition on $X_1 = x$, then the other $X_2, ..., X_N$ are still random. We would have to condition N times (a random number of conditions!) to make that work.

Lecture 10 - Jan 25

[Diana Katherine Skrzydlo] - Jan 27, 2012 8:23 AM
Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

Today we started talking about computing variances by conditioning. We have a couple of techniques we can use.

We can find a recursive equation for $E[X^2]$ by finding $E[X^2|Y=y]$ for every possible value of y and then using double averaging. We did an example of this with the same mouse in the maze. Remember we square the entire variable X|Y=y before taking its expected value.

Or, we can use the conditional variance formula. This works great if the distribution of X|Y=y is something we recognize. We defined Var(X|Y=y) and the corresponding random variable Var(X|Y) similar to the expected value case, and then proved the formula Var(X) = E[Var(X|Y)] + Var(E[X|Y]). We did an example of this technique as well.

Lecture 9 - Jan 23

Diana Katherine Skrzydlo - Jan 23, 2012 9:48 AM
Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

Today we did an example of two independent random variables and using the properties of conditional expectation to find the distribution of a combination of them. I did it by finding the tail probability P(Z>z) but you should theoretically get the same result if you do P(Z<= z) directly. You will end up with two integrals (one from 0 to min{1/z, 1} and one from min{1/z, 1} to 1).

Then we did a generalization of a logic problem - n offers in n envelopes, and if you open an offer you must immediately accept or reject it. It turns out the optimal strategy for finding the best offer is to open k envelopes and then choose the first one that is larger than all those. The k that maximizes the probability of getting the best offer is n/e. (I didn't show the maximizing, but you can check out the excel sheet which graphs the probability for various values of n and k). Apparently a radio station is doing this for a contest with 5 envelopes. 5/e rounds to 2, just so you know if you get on the show!

Next time we'll talk about computing variances by conditioning.

Lecture 8 - Jan 20

Diana Katherine Skrzydlo - Jan 20, 2012 9:50 AM
Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

Today we finished the example of two continuous rvs.

We looked at 4 properties of conditional expectation (a handout is posted) and how we can apply them in 3 cases based on the information we're given. We also described how these techniques can be used for probabilities as well as expectations, since a probability of an event is just an expectation of an indicator rv attached to that event.

We did an example of a discrete X continuous Y (we called it U) using our knowledge of the conditional distribution of X|U=u (this was case 2)

Lecture 7 - Jan 18

Diana Katherine Skrzydlo - Jan 20, 2012 9:47 AM
Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

Today we used conditioning to solve a problem where the distribution of X on its own is quite complicated using the setup of the mouse in the maze. If we condition on the door the mouse chooses first, it becomes straightforward.

We also extended our definitions to the continuous case (everything stays pretty much the same except for the interpretation) and started an example of two continuous random variables starting from their joint pdf.

Lecture 6 - Jan 16

Diana Katherine Skrzydlo - Jan 16, 2012 4:38 PM
Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

Today we started Chapter 3 with a whole lot of definitions. We have the conditional rv X|Y=y, its conditional range, conditional pmf P(X=x|Y=y), and conditional expectation E[X|Y=y]. Then if we allow y to vary as well, we have a new random variable E[X|Y] which is just a function of Y. We proved (in the discrete case) that E[E[X|Y]] = E[X] for any rv Y, which will be very useful for lots of problems. We saw one example which we will continue next time.

Note: P(X=x|Y=y) is a valid pmf if X is proper to begin with and P(Y=y) is not 0. Similarly, I didn't mention this, but E[X|Y=y] is guaranteed to converge as long as X is *short* proper.

Lecture 5 - Jan 13

Diana Katherine Skrzydlo - Jan 13, 2012 12:13 PM
Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

Today we talked about indicator variables and how they can be used to solve complicated problems. We did an example involving the number of unrolled faces on a die after n trials. I'll provide a graph of what the mean and variance look like for various values of n and post a link here when it's up.

We also looked more closely at the classification of random variables into improper, short proper, and null proper, and the conditions for each.

Have a great weekend!

Lecture 4 - Jan 11

Diana Katherine Skrzydlo - Jan 13, 2012 8:20 AM
Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

Today we talked about continuous random variables, summarized the distributions we care about in this course in a chart, and briefly went over some properties of expectations of (functions of) random variables, including variance. We also talked about joint distributions, covariance, and how to find out if two random variables are independent.

Questions from after lecture
How come Poisson was under the "continuous" category if it's a discrete random variable?
The headings I gave are the situation that leads to the distribution. In order to have a Poisson random variable (which is indeed discrete, remember it's the range of the rv that determines the type), we need a continuous process indexed by time or something similar, where events can occur at any instant. In that context, if we count up the number of events in a fixed time interval, we get a Poisson distribution.

If a joint pmf or pdf f(x,y) factors into two functions g(x)*h(y), does that automatically mean they are independent, and that g(x) is the pdf of X and h(x) is the pdf of Y?
Yes, but ONLY if the ranges of X and Y are also independent of each other. Or equivalently, ONLY if the range of (X,Y) is a product of sets. For example, if the range is $0 < x < y$, or $x > 0$, $y > 0$, $x+y < 1$, those are not independent, no matter what the pdf factors into.

Lecture 3 - Jan 9

[Diana Katherine Skrzydlo]- Jan 9, 2012 5:22 PM
Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

Today we talked all about random variables. We defined them, classified them according to their range (not domain), and defined functions that describe them (pmf in the discrete case, pdf in the continuous case, cdf in both).

We looked at Bernoulli trials as a way to give us several distributions: bernoulli/indicator, binomial, geometric, and negative binomial, and looked at the relationships between them and the property that the geometric distribution is proper as long as p > 0. That means no matter how unlikely something is, as long as it's not impossible, it is guaranteed to occur in a finite number of trials (for example, an infinite number of monkeys on an infinite number of typewriters will eventually produce the complete works of Shakespeare.)

We also saw the Poisson random variable, and we'll look at a few continuous ones next time.

Lecture 2 - Jan 6

Diana Katherine Skrzydlo - Jan 6, 2012 10:15 AM
Mark Unread
[Reply]
More actions...

0 unread of 5 messages - 3 author(s)

Today we finished off Chapter 1. We talked about combinations of sets, and defined a probability as a function mapping subsets of S to real numbers, subject to 3 axioms. We can prove all the properties of probability with those axioms. A summary of the material is posted in Content under Review Quiz. We also defined conditional probability and Bayes' rule.

The Monty Hall problem is an application of Bayes' rule. I'll get it started - see if you can figure out the rest.
Let's let the events A, B, and C be "the car is behind door A, B, or C" respectively. Then we know P(A) = P(B) = P(C) = 1/3.
Let's let the event O be the event "Monty opens a door, revealing a goat"
Without loss of generality, we can suppose you choose door A and he opens door B.
Now what we're looking for is the probability you win if you switch, which is the probability the car is behind door C. BUT we have the conditional information that he opened B, so we are actually looking for P(C|O).
Try using Bayes' rule on this probability and see if you get 2/3.

Question about immeasurable sets

A probability must be defined on a measure space, and thus we restrict our definition of probability to measurable subsets of S. Great question, I had to look up my STAT 901 notes to find out! Incidentally, if you want more info about the Cantor set, google it. I'll post if I find a good page.

Collapse Replies

Re: Lecture 2 - Jan 6 (Monty Hall)

Diana Katherine Skrzydlo - Jan 9, 2012 8:13 AM
Mark Unread
[Reply]
More actions...

0 unread of 4 messages - 3 author(s)

What we're looking for is $P(C|O)$.
Using Bayes' rule, we can express this as $P(O|C)*P(C)/P(O)$
But $P(O)$ is pretty difficult to calculate directly, so we'll use the law of total probability on it.

The event O can be split up into three disjoint parts: $O = OA \cup OB \cup OC$ and we can apply the multiplication rule to each part to get $P(O) = P(A)*P(O|A) + P(B)*P(O|B) + P(C)*P(O|C)$
Obviously $P(O|B)$ is 0 (since he cannot open door B to reveal a goat if B has the car.
$P(O|C)$ is 1 (since if you chose A and the car is behind C, the only door he can open is B)
$P(O|A)$ is 1/2 (since he randomly chooses a goat-door to open)

Putting this all together,
$P(C|O) = P(O|C)*P(C)/P(O)$
$= P(O|C)*P(C)/[P(A)*P(O|A) + P(B)*P(O|B) + P(C)*P(O|C)]$
$= 1*(1/3) / [(1/2)*(1/3) + 0*(1/3) + 1*(1/3)]$
$= (1/3) / (3/6)$
$= 2/3$ as required.

Neat, eh? Can you think of a simpler, logical explanation for this non-intuitive result?

Collapse Replies

Re: Lecture 2 - Jan 6 (Monty Hall) 0 unread of 1 messages - 1 author(s)

[Student](#) - Jan 9, 2012 6:31 PM
[Mark Unread](#)
[[Reply]](#)
[More actions...](#)

Is this correct?

Not swapping:

Originally, you have 33% of picking the car. Your strategy is not to switch so you stay 33% (thus, getting a goat is 67%).

Swapping:

Originally, you have 33% of picking the car. Your strategy is to switch every time. If your door has a car, then you switch to receive the goat--a 33% occurrence. However, to pick a goat and switch (to the only available choice: a car) is 67% likely.

I think getting that one door open effectively switches the probabilities of choosing a goat and choosing a car. If there were 1000 doors, you choose 1 (.1% chance that you guessed the car) but then 998 other doors open revealing all goats. What's the likelihood of the last door having the car and not a goat?

On second thought, this question may have been rhetorical. D'oh.

<<< Replied to message below >>>
Authored by: Diana Katherine Skrzydlo
Authored on: Jan 9, 2012 8:13 AM
Subject: Re: Lecture 2 - Jan 6 (Monty Hall)

What we're looking for is $P(C|O)$.
Using Bayes' rule, we can express this as $P(O|C)*P(C)/P(O)$
But $P(O)$ is pretty difficult to calculate directly, so we'll use the law of total probability on it.

The event O can be split up into three disjoint parts: $O = OA \cup OB \cup OC$ and we can apply the multiplication rule to each part to get $P(O) = P(A)*P(O|A) + P(B)*P(O|B) + P(C)*P(O|C)$
Obviously $P(O|B)$ is 0 (since he cannot open door B to reveal a goat if B has the car.
$P(O|C)$ is 1 (since if you chose A and the car is behind C, the only door he can open is B)

P(O|A) is 1/2 (since he randomly chooses a goat-door to open)

Putting this all together,
P(C|O) = P(O|C)*P(C)/P(O)
= P(O|C)*P(C)/[P(A)*P(O|A) + P(B)*P(O|B) + P(C)*P(O|C)]
= 1*(1/3) / [(1/2)*(1/3) + 0*(1/3) + 1*(1/3)]
= (1/3) / (3/6)
= 2/3 as required.

Neat, eh? Can you think of a simpler, logical explanation for this non-intuitive result?

Re: Lecture 2 - Jan 6 (Monty Hall)

Student - Jan 11, 2012 4:25 AM
Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

A simpler explanation would be to regard this as the change of variable from a car to goat.

Switching loses if and only if the player initially picks the car, which has a probability of 1/3, so switching must win with probability of 2/3.

Suppose the player initially picked a goal, he will definitely win a car by switching since the other goat can no longer be picked. The probability of the player initially selecting a goat is 2/3.

Re: Lecture 2 - Jan 6 (Monty Hall)

Diana Katherine Skrzydlo - Jan 13, 2012 8:13 AM
Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

Right, both very good alternative explanations. There are lots!

Lecture 1 - Jan 4

Diana Katherine Skrzydlo - Jan 4, 2012 9:56 AM
Mark Unread

0 unread of 2 messages - 2 author(s)

[Reply]
More actions...

Today we went over the course outline.

We started by brainstorming material from STAT 230 and came up with a great list of concepts that we'll use in this course. Then we began Chapter 1 by defining experiment, sample space, trial, sample point, and event. We can look at combinations of events using the union, intersection, and complement operators (which we will continue next time.)

A couple of questions from after class

- What is the difference between countable and uncountable infinity?
Countably infinite means we can map it to the integers, i.e. put them into some order and count them. Any finite set, the positive integers, the entire set of integers, and even the rational numbers are countable sets. Uncountable means we cannot make a list of all the items in the set - if we try, we will leave some out.

- Can we define different sample spaces on the same experiment?
Yes. In our example where the experiment was "roll a die", the obvious sample space is S = {1, 2, 3, 4, 5, 6}. But S = {even, odd}, S = {prime, not prime}, S = {1, 3, 5, even}, etc are all valid sample spaces since they include all the possibilities and nothing is double-counted. They're not as useful though, since any events we are interested in have to be subsets of S.

- If S is continuous, is the power set of S an event?
I'll think about this one and get back to you here on this blog

Collapse Replies

Re: Lecture 1 - Jan 4

Student - Jan 4, 2012 6:20 PM
Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

In my opinion, if S is the sample space (no matter countable or not), then the power set of S is the set containing all the events (subsets of S) of the underlying experiment, and it is not an event itself because the power set is not a subset of S but a set of events.

Actually there is an elegant way to tell the power set of a set A is never a subset (or event in this case) of A by cardinality rather than repeating "sets" and "events" like above. The power set of A always has a higher cardinality than A (by Russell's Paradox), which implies that the power set of A can never be a subset of A, otherwise the power set of A would has a cardinality no greater than does A.

However, under the subset definition of events, it is possible to have a set of outcomes identical to another outcome and a set of events identical to another event, for example S={a,{a}} where a could be anything, but such confusions could be eliminated by changing notations or the definition of events if one really wants to try. The following example is just for some fun. Set theorists define all natural numbers with only the empty set and set operations. In that way, 0=empty and 1={empty}. So a sample space like {0,1} would be {empty,{empty}}, which leads to confusion as mentioned above. But we are all OK with a sample space like that since we all think of 1 as an outcome instead of an event of observing 0.

Anyway, we are focusing on probability theory rather than set theory in this course: figuring that out may not be helpful. Most importantly, life is meant to be easy and we seldom, if any, encounter weird things like that.

- Kevin

<<< Replied to message below >>>
Authored by: Diana Katherine Skrzydlo
Authored on: Jan 4, 2012 9:56 AM
Subject: Lecture 1 - Jan 4

Today we went over the course outline.

We started by brainstorming material from STAT 230 and came up with a great list of concepts that we'll use in this course. Then we began Chapter 1 by defining experiment, sample space, trial, sample point, and event. We can look at combinations of events using the union, intersection, and complement operators (which we will continue next time.)

A couple of questions from after class

- What is the difference between countable and uncountable infinity?
Countably infinite means we can map it to the integers, i.e. put them into some order and count them. Any finite set, the positive integers, the entire set of integers, and even the rational numbers are countable sets. Uncountable means we cannot make a list of all the items in the set - if we try, we will leave some out.

- Can we define different sample spaces on the same experiment?
Yes. In our example where the experiment was "roll a die", the obvious sample space is S = {1, 2, 3, 4, 5, 6}. But S = {even, odd}, S = {prime, not prime}, S = {1, 3, 5, even}, etc are all valid sample spaces since they include all the possibilities and nothing is double-counted. They're not as useful though, since any events we are interested in

> have to be subsets of S.
>
> - If S is continuous, is the power set of S an event?
> I'll think about this one and get back to you here on this blog