ACTSC 613 Lecture Blog Fall 2012

Lecture Blog

I will use this space to post brief summaries of the lectures shortly afterwards. You can use it to see what you missed if you couldn't be at a class (though you should also get notes from a classmate) and also for review of the main concepts.

Lecture 23 - Dec 3

Diana Katherine Skrzydlo - Dec 3, 2012 5:46 PM

Mark Unread [Reply] More actions...

0 unread of 1 messages - 1 author(s)

Today we talked about multiple linear regression. We use the same ideas as for a simple model, but now we can express any model with p regressors in vector and matrix form as $\underline{y} = X\underline{\beta} + \underline{\varepsilon}$.

We talked about interpretation of the parameters in this context, and residual analysis.

To estimate the parameters, we use least squares estimation to get the p+1 normal equations $(X^TX)\underline{\beta} = X^T\underline{y}$, which we then solve for the β-hat vector $(X^TX)^{-1}X^T\underline{y}$. Once we have those, we can derive the fitted values and residuals as well. All of them are just linear combinations of the yi's. Since the yi's are normal (if the εi's are, which we assume), the β-hats, fitted values, and residuals are all normal as well, and we can derive their means and variances (and covariances). Then we can use their distributions to construct confidence intervals or do hypothesis tests for individual β_j's as in the simple regression case.

Lastly, we looked at how the ANOVA changes with a multiple linear regression. The good news - not very much! We just have p degrees of freedom for the SSR, n-p-1 for the SSE, and all the other relationships in the table are the same. The F-test in the table is testing H_0: $\beta_1 = \beta_2 = \ldots = \beta_p = 0$ (all 0 at the same time.)

Lecture 22 - Nov 28

Diana Katherine Skrzydlo - Nov 28, 2012 11:57 PM

Mark Unread

0 unread of 1 messages - 1 author(s)

[Reply]
More actions...

Today we talked about Linear Regression, which is a method to determine the relationship between one (or more) explanatory variables (x) and a response variable (y).

For today's class we focused on the simplest model with one x, and only a slope and intercept parameter. So $y = \beta_0 + \beta_1 x + \varepsilon$ is our model. Next time we'll see the general case. We assume the ε's are iid Normal with mean 0 and variance σ^2.

We used least square estimation to derive the normal equations and hence the least squares estimates of β_0 and β_1. Then we can calculate the fitted values and residuals from the model. The residuals can be used to check the model for adequacy, as well as to estimate σ^2.

We can derive CIs and hypothesis tests for the unknown parameters and for the average (or a single new) value of the response for a particular x value x_p. All of those calculations require the original assumption that the ε's were Normal and independent.

Finally, we looked at ANOVA for a simple regression model. We can split the variability in y (SST) into the part that is explained by the model (SSR) and the part that is unexplained (SSE). The R^2 value, SSR/SST, represents the proportion of variability explained by the model, and we can also build an ANOVA table by using the degrees of freedom for each SS. The table includes an F-test which is equivalent to the hypothesis test for $\beta_1 = 0$ in this case.

Throughout, we used R (both manually calculating and with the assistance of the "lm" and "anova" functions) to demonstrate the techniques. The code we used is posted along with some comments.

Lecture 21 - Nov 26

Diana Katherine Skrzydlo - Nov 26, 2012 5:44 PM
Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

We reviewed the procedure for testing the mean in a case where the variance is unknown, and talked about testing the variance, testing both the equality of means and equality of variances in a two-sample problem, and testing the mean in a paired-sample problem. The general idea is always the same.

Then we looked at the Chi-Squared goodness of fit test, which is a neat procedure we can use to test whether data follows a particular distribution (be it binomial with a particular p_0, multinomial, or otherwise) whether we have one or more samples from the distribution. It always has the form of calculating the sum of $(o_j - e_j)^2/e_j$ over all cells, and comparing to a

Chi-Squared distribution with df = # of cells - # of restrictions.

We discussed the various uses of the Chi-Squared test, including single sample multinomial, multiple sample binomial and multinomial, testing for data following a particular distribution, and testing for independence between two factors (called a contingency table.) We saw a couple of examples of the tests.

In tutorial we looked at problem #4 from Ch 12, and #11 and #13a from Ch 13.

I also was going to have you do #3 from Ch 11, so I'll post the solution to that here.

Lecture 20 - Nov 21

Diana Katherine Skrzydlo - Nov 26, 2012 10:56 AM

Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

Today we started Hypothesis Tests, which is a more formal way to draw conclusions from data. We defined the null and alternative hypothesis, the simple and composite hypothesis pair, and the two types of error as well as their probabilities (and what they are called in the simple and composite case). We looked repeatedly at an example of finding a critical region for testing whether the mean of a normal distribution was 10 under various different alternative hypotheses, and found the power of the test in one or two cases. We also found the p-value in one case, which is the probability of observing a value as extreme (or worse) as we saw, if the null hypothesis were true.

We talked about the difference between one-sided and two-sided tests, and looked at the procedure for testing the mean in a case where the variance is unknown. The general idea is always the same, as we will see next time with more examples.

Lecture 19 - Nov 19

Diana Katherine Skrzydlo - Nov 19, 2012 5:41 PM

Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

Today we finished up chapter 11. We talked about the 2-sample problem and finding a CI for the difference in means when the variances are unknown. We can assume that the unknown variances are equal and then get a pooled estimate of the variance, and then derive a CI for $\mu_2 - \mu_1$. If the variances are not equal, unfortunately it is very difficult (and I won't be talking about that case.)

We did an example of finding CIs for both the ratio of variances and difference of means using some data.

Finally we talked about the paired-sample or paired-comparison problem, which is when we have two observations (X_i and Y_i) on each of the experimental units, i = 1, ..., n. In that case, we should look at the differences for each unit ($D_i = Y_i - X_i$) and analyze that like a one-sample problem.

Lecture 18 - Nov 14

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

0 unread of 1 messages - 1 author(s)

Today we started chapter 11, confidence intervals.

We defined a CI (in terms of the random interval (L, U) and the actual interval (l(x), u(x))) and its interpretation informally and formally. We defined a pivotal quantity, and saw examples of finding a CI for the mean of a Normal distribution with known variance, the mean and variance of a normal distribution where both are unknown, and how we can use the asymptotic distribution to find approximate CIs for other distributions such as Bernoulli and Poisson.

Finally we started looking at the two-sample problem where we are interested in comparing the variances and means of two populations. We derived the CI for the ratio of the variances, and for the difference in the means, if we assume the variances are known. Next time we'll see what happens if the variances are unknown.

Lecture 17 - Nov 12

Diana Katherine Skrzydlo - Nov 12, 2012 6:02 PM
Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

Today we finished off chapter 9 by talking about general loss/risk functions (including the MSE, mean squared error, which takes into account both the variance and the bias of an estimator).

We also looked at Bayesian estimation, a really neat method which allows us to make use of our knowledge of the distribution before any data is collected (the prior distribution), as well as the information gained from our sample. We get the posterior distribution of the unknown parameter Θ given the sample $X_1, ..., X_n$, by applying Bayes' rule to reverse the conditioning, and then the Bayesian estimator of θ is just the mean of that distribution.

We saw an example in detail and I'll write up the results from the other one and post them.

In tutorial we looked at questions 1, 3, and 28 from Chapter 9.

Lecture 16 - Nov 7

Diana Katherine Skrzydlo - Nov 7, 2012 5:28 PM
Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

Today we described the Method of Moments (where we equate the theoretical moments of the distribution with the sample moments of the data) and did two examples.

We then talked about maximum likelihood estimation, including defining the likelihood and log-likelihood functions to maximize with respect to the unknown parameter over the parameter space. We did three examples, the last of which has an interesting practical application estimating the number of tanks produced in WW2. (http://www.guardian.co.uk/world/2006/jul/20/secondworldwar.tvandradio).

Lecture 15 - Nov 5

Diana Katherine Skrzydlo - Nov 5, 2012 4:18 PM
Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

Today we finished off chapter 8 by looking at three ways we can draw conclusions about unknown parameters from a normal sample.
- $(\bar{X} - \mu)/(S/\sqrt{n})$ has a T distribution with n-1 degrees of freedom so we can examine the true unknown mean
- $(n-1)S^2/\sigma^2$ has a Chi-squared distribution with n-1 degrees of freedom so we can examine the population variance
- For two samples, $(\sigma_2^2 S_1^2)/(\sigma_1^2 S_2^2)$ has an F distribution with $n_1 - 1$ and $n_2 - 1$ degrees of freedom so we can examine the ratio of the population variances

Then we started talking about parameter estimation methods. We considered several properties that are desirable for an estimator to have, including being invariant, unbiased, UMVUE (aka efficient), consistent, asymptotically unbiased, and asymptotically minimum variance (and also a way to compare two estimators using their relative efficiency or their asymptotic relative efficiency.)

Lecture 14 - Nov 2

Diana Katherine Skrzydlo - Nov 2, 2012 1:14 PM

Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

(Oct 29 was midterm review, Oct 31 was the midterm)

Today we started the "statistics" part of the course (rather than the "probability" part)

We defined a statistic, which is a function of a random sample (so it's a random variable) that does not contain any unknown parameters, as well as an estimator and an estimate. We examined the sample mean and the sample variance, showing both are unbiased estimators. We reviewed the chi-squared distribution, which we used to prove the main result for a random sample of size n from a $N(\mu, \sigma^2)$:
1) the sample mean X-bar is independent of the differences from the sample mean (X_i - X-bar)
2) the sample mean X-bar is independent of the sample variance S^2, and
3) $(n-1)S^2/\sigma^2$ has a chi-squared(n-1) distribution

Lecture 13 - Oct 24

Diana Katherine Skrzydlo - Oct 24, 2012 6:13 PM

Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

Today we finished up the probability half of the course!

We talked about stochastic convergence (convergence to a constant, which is the case where both convergence in distribution and in probability are the same thing), and looked at an example. We also defined convergence in mgf and showed (another way to using the pmf) that the limiting distribution of the binomial is the poisson.

We proved the Central Limit Theorem and showed how we can use it to approximate sums of random variables (or even some random variables on their own) when the n is reasonably large. We also looked at how to improve the approximation by using a continuity correction when using the CLT on a discrete distribution.

Finally, we proved the law of large numbers (which has lots of great implications for diversifiable risk in the life insurance context, among other things.)

Lecture 12 - Oct 22

Diana Katherine Skrzydlo - Oct 22, 2012 5:53 PM
Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

Today we talked about conditional variance (including the conditional variance formula) and compound random variables.

We also looked (theoretically) at joint transformations. We'll see a very important one of these later on in the term, but it's a useful extension to n random variables of the pdf technique.

Finally, we started chapter 7 by talking about two kinds of convergence (there are others) - convergence in distribution and convergence in probability. We'll see more kinds and more results and examples in Wednesday's class.

In the tutorial we had a sort of review for the Midterm (though understandably not many of you have started studying for it yet). If you have any questions about the problems from the textbook, past test papers, etc, please feel free to post on the discussion board or come see me. If you don't want to post publicly, you can email me your question and I'll copy/paste it (without saying who asked it) and answer it on the discussion board. But don't be shy, it's pretty likely that someone else has the same question that you have!

Lecture 11 - Oct 17

Diana Katherine Skrzydlo - Oct 17, 2012 11:12 PM
Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

Today we wrapped up a few extra topics about joint random variables (correlation and using generating functions of multiple independent random variables).

Then we talked about all things conditional. We defined conditional random variables (say X|Y=y), their conditional pmfs (and later pdfs where the interpretation is slightly different)

$f(x|y)$, conditional expectations $E[X|Y=y]$. We also defined a new random variable $E[X|Y]$ which is actually a function of the random variable Y, and looked at some of the properties of conditional expectation. We can use the properties to calculate expected values as well as probabilities by conditioning. Next time we'll see how to use conditioning to calculate variances as well.

All the assignment 2's are now marked and I'll put the 4 remaining (sorry to you 4!) outside my office in the folder tomorrow morning for you to pick up. I'll also distribute my feedback on the presentations from assignment 2 individually next week.

Lecture 10 - Oct 15

Diana Katherine Skrzydlo - Oct 15, 2012 5:48 PM
Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

Today we talked about continuous joint distributions (joint cdf, joint pdf, marginal pdfs, and their properties), and what it means for variables to be independent. We looked at expectations of functions of multiple random variables and the linearity property. In particular, we can look at the covariance between two random variables, which tells us whether they have any linear relationship.

In the tutorial we looked at Ch 4 #8, 9, 14, and Ch 5 #4 and 11.

Lecture 9 - Oct 10

Diana Katherine Skrzydlo - Oct 11, 2012 11:23 AM
Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

Today we started chapter 4 - joint random variables.

We looked at the discrete case, defining the joint pmf, marginal pmfs, joint cdf, and looking at their properties. We saw two examples - a simple one where we define the joint pmf by a chart (and can get all the other information from it), and the multinomial distribution which is an extension of the binomial distribution.

We had the second set of presentations in the tutorial. I posted sketches of the 5 proofs I can think of that NB is proper, in case you're interested. There may be other valid proofs too!

Next time we'll see the continuous case and look at conditional distributions.

Lecture 8 - Oct 3

Diana Katherine Skrzydlo - Oct 3, 2012 5:33 PM
Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

Today we started with a giant chart of all the distributions and the relationships between them. There's a neat online clickable version of something similar link posted in Content.

Then we talked about generating functions, including probability (or factorial moment) generating functions, moment generating functions, and cumulant generating functions. We'll be using them a bit more in the future, and they are especially useful for proving properties and relationships for distributions.

There is no class on Monday, so see you in a week! Please feel free to post on the discussion board or email me if you have more questions about the assignment.

Lecture 7 - Oct 1

Diana Katherine Skrzydlo - Oct 2, 2012 5:13 PM
Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

Today we talked about the rest of the continuous random variables: Weibull, Pareto, Cauchy, and some sampling distributions (details in Ch 8) being the Chi-squared, t, F, and Beta. I posted a picture which does a better job than my little stuffed toy distributions comparing the various t distributions to a standard Normal.

Next time we'll summarize all the distributions (discrete and continuous) and their relationships in a way that will hopefully help remember them all! And we'll also look at generating functions.

We also had the second tutorial, and looked at questions 1, 4, 11, 16, and 21 in Chapter 3. I posted a spreadsheet for question 16 that lets you play around with the probability of A winning an individual game to see what happens.

Lecture 6 - Sept 26

0 unread of 1 messages - 1


[Diana Katherine Skrzydlo](#) - Sep 28, 2012 12:00 AM
[Mark Unread](#)
[Reply]
[More actions...](#)

author(s)

Today we talked about the Poisson random variable (along with the two contexts in which it arises - the limit of the Binomial and the Poisson Process)

We also looked at the first few continuous distributions - uniform, exponential, gamma, normal, and lognormal. We also talked about transformation techniques (specifically the cdf and pdf techniques) and the relationships between some of the distributions.

Someone asked afterwards why we know that $\lim_{n\to\infty} (1 + k/n)^n = e^k$. This is actually one of the equivalent definitions of the constant e itself (see http://en.wikipedia.org/wiki/E_(mathematical_constant)) for more details.

Lecture 5 - Sept 24


[Diana Katherine Skrzydlo](#) - Sep 24, 2012 5:13 PM
[Mark Unread](#)
[Reply]
[More actions...](#)

0 unread of 1 messages - 1 author(s)

Today we started talking about some of the specific distributions we will be interested in in this course. We covered the discrete uniform, bernoulli/indicator, binomial, hypergeometric, geometric, and negative binomial, including their definitions, ranges, pmfs, means, and variances. None of those are things I would expect you to memorize, but it's good to be familiar with them. We also looked at some of the relationships between the distributions - we'll continue to see more relationships as we go.

We also had the first presentations, which were quite good! I'll be giving you individual feedback along with your marked assignments as soon as I can.

Lecture 4 - Sept 19


[Diana Katherine Skrzydlo](#) - Sep 19, 2012 11:21 PM
[Mark Unread](#)
[Reply]
[More actions...](#)

0 unread of 1 messages - 1 author(s)

Today we finished chapter 2.

We defined the expected value of a random variable and looked at some of the other measures of a random variable (median, mode, and percentiles) that can provide summary information. Expectation is the most important, and we looked also at expected values of a function of a random variable, including particularly the variance and other moments (about the origin or about the mean).

We ended by looking at some ways to obtain bounds on probabilities, if all we have is the mean/variance of functions of a random variable rather than its exact distribution.

Lecture 3 - Sept 17

Diana Katherine Skrzydlo - Sep 19, 2012 11:18 PM

Mark Unread

[Reply]

More actions...

0 unread of 1 messages - 1 author(s)

Today we started chapter 2 by talking about random variables.

We define a random variable X as a function that maps points in the sample space to real numbers. We can then express the probability that X takes on a particular set of values A as the probability of the event given by the set of sample points that X maps to A, and then apply the probability function P. Random variables are classified as discrete or continuous based on the range (not the domain) of the function X.

For discrete random variables, we can define the probability mass function f(x) = P(X=x) and the cumulative distribution function F(x) = P(X <= x). They have various properties which follow from the axioms of probability.

For continuous random variables, the cdf is the same but we define the probability density function f(x) = F'(x) where the derivative exists. f(x) is not a probability, but represents the amount by which we multiply ε to obtain the approximate probability that X is within ε of the value x.

We can also have mixed random variables which have both continuous and discrete pieces.

We then had the first tutorial where we looked at questions 3, 19, and 29 in chapter 1 and questions 3 and 17 in chapter 2.

Lecture 2 - Sept 12

0 unread of 2 messages - 2 author(s)

Diana Katherine Skrzydlo - Sep 12, 2012 5:17 PM
Mark Unread
[Reply]
More actions...

Today we finished off chapter 1.

We talked about conditional probability, the definition, formula, and justification. We can rearrange it to get the multiplication rule, and use both together to get Bayes' rule which allows us to reverse the conditioning on a pair of events. We can also use partitioning and the multiplication rule to find the unconditional probability of any event, which is called the law of total probability. We looked at two examples (one using a tree diagram as a tool) including the Monty Hall problem. Finally, we talked about independent events.

The counting techniques in section 1.6 is not something I'm going to specifically teach or test, but you should be familiar with "choose" and "factorial" concepts.

Next time we'll start chapter 2 and talk about random variables, the functions associated with them and their properties.

Collapse Replies

Re: Lecture 2 - Sept 12

Ming Wu - Sep 13, 2012 11:05 PM
Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

Haha, the Monty Hall problem is really interesting~

<<< Replied to message below >>>
Authored by: Diana Katherine Skrzydlo
Authored on: Sep 12, 2012 5:17 PM
Subject: Lecture 2 - Sept 12

Today we finished off chapter 1.

We talked about conditional probability, the definition, formula, and justification. We can rearrange it to get the multiplication rule, and use both together to get Bayes' rule which allows us to reverse the conditioning on a pair of events. We can also use partitioning and the multiplication rule to find the unconditional probability of any event, which is called the law of total probability. We looked at two examples (one

using a tree diagram as a tool) including the Monty Hall problem. Finally, we talked about independent events.

The counting techniques in section 1.6 is not something I'm going to specifically teach or test, but you should be familiar with "choose" and "factorial" concepts.

Next time we'll start chapter 2 and talk about random variables, the functions associated with them and their properties.

Lecture 1 - Sept 10

Diana Katherine Skrzydlo - Sep 10, 2012 5:00 PM
Mark Unread
[Reply]
More actions...

0 unread of 1 messages - 1 author(s)

Today we discussed the course syllabus.

Then we started Probability by defining several terms: experiment, trial, outcome, sample space (discrete vs continuous), and event. We talked about how to combine events with the set operators U (union), ∩ (intersection), and ' (complement), simple vs compound events, and events that are equivalent, mutually exclusive, and a partition.

We defined probability formally as a function that maps events (subsets of S) to Real numbers according to 3 axioms: the probability of any event is non-negative; the probability of the entire sample space is 1; and for any number of mutually exclusive events, the probability of the union is equal to the sum of the probabilities. We used the axioms to prove some relatively straightforward properties of probability.

Next time we'll talk about conditional probability, Bayes' rule, and independent events.
