Models In Epidemiology And Biostatistics
Gordon Hilton Fick

Session 13 :
Linear Regression – Studying Conditional Means

Up to now, we have been exploring the range of methods for studying an outcome that takes on a short list of possibilities. Indeed, we were concerned with the simplest possible outcome: Yes or No, Success or Failure, Alive or Dead... the outcome with only two possible values.[coded 0 and 1]

We have addressed outcomes with more than two values. The implication [maybe not said] was that the list of possible values is short enough so that direct study of each of these possible values has merit. All our approaches involved rates or odds; ratios or differences to make statements about each of the possible values of the outcome.

Continuity:
Let us now focus our attention on outcomes with a potentially long list of values. It may be valuable to think about outcomes that are said to be absolutely continuous in that there are so many possible values that is becomes impossible to speak of the probability that an outcome takes on an exact value. A person's exact age [in days, say] is the easiest example. A person's systolic blood pressure (sbp) [mmHg] is another. Of course, the measurement of either age or sbp is done with finite accuracy and inevitably involves round off or grouping into intervals. Even so, there are typically so many such intervals that direct study of all possible intervals is out of the question.

A return to the methods of earlier classes:
Often a credible option for analysis to to decide on thresholds or intervals that determine a new outcome that has a small list of values. For example, one might wish to consider models for "sbp >130 mmHg" rather than attempt to directly study sbp. Another example from diabetes research involves a characteristic called AER [albumin excretion rate: an indicator for kidney trouble]. Researchers can study AER directly but will often use 2 thresholds to yield 3 intervals: AER<20 (normal); 20<AER<200 (microalbuminuria) and AER>200 (macroalbuminuria). Often such a construction of a 'new' outcome is judged to be more 'practical' or 'more clinically relevant'. On the other hand, specialists argue about the merits of such thresholds: their cutoff values and the number of such cutoffs. The decision to adopt cutoffs/thresholds/intervals does not necessarily lead to inferior assessment compared to the methods of this chapter. There are some writers who might argue that it is always to advantage to use the actual values rather than to use intervals. Such uncategorical statements serve little purpose. Alas, it would seem that most of the time, it depends....

The conditional distribution of the outcome.
Now we imagine a collection of possible exposures and confounders/modifiers and the decision has been made to attempt to study the distribution of the outcome given [or conditional on] such exposures/confounders/modifiers. We will, for the time being, suppose that these conditional distributions are all symmetrical. With this symmetry, we can usually speak meaningfully about the mean of such a distribution in so far as such a mean reflects the 'centre' of the distribution and is the same as the median and [for unimodal distributions] the mode. Accordingly, for the time being, we will focus on the conditional mean [sometimes called the conditional expectation and also called the regression]

Case Control Studies
Notice that for case-control studies, we may wish to directly study an absolutely continuous exposure variable as a primary outcome. For example [from the endometrial cancer study], we may wish to study

the length of time a woman was receiving estrogens. For this type of study, we then wish to study the distribution of exposure given [or conditional on] case/control status, confounders and modifiers.

Symmetry Assessment
As we shall see, the assumption of symmetry is typically critical to meaningful interpretation. We will also see that the assessment of the assumption of symmetry must wait until we can understand the nature of the distributional form as though there were no dependency of the conditional mean on the respective conditions. [Hard to appreciate just now... sometimes a rereading can help... witness the phenomenon of statistics wannabees reading and rereading RA Fisher over and over and over....]

...so we launch into the BIG topic of Linear Regression. We will see that primary to this topic is clearly appreciating that we are always making conditional statements: statements about a conditional mean and the when and how such means depend on the condition(s). We will soon see that the 'distribution form' of variables on the 'right hand side of the regression equation' is not relevant.

Linear regression means that the conditional mean is linear in the predictor variables

$$E(y)=\sum_{j=0}^{k}\beta_j x_j$$

As we did with logistic regression, we won't explicitly list the conditions in the expectation (left hand side). Some might say that it is clearer to write:

$$E(y|x_0\ \ x_1\ \ x_2\ \ ....\ \ x_k)=\sum_{j=0}^{k}\beta_i x_i$$

We will not take this path. For us, a mean is ALWAYS a conditional mean just as odds ratio is always a conditional odds ratio. The list of conditions [the $x_i's$] are the exposures, the confounders, the modifiers and other explanatory variables. As previously, we need to make clear the distinction between linearity, additivity, interaction, confounding, modification and the complicated forms of confounding and modification. All of these issues carry forward to linear regression. Indeed, in so far as we are interpreting the coefficients, we now need to speak of means or estimates of means rather than speaking of log of odds and/or estimates of log of odds [or the others...]. The examples will help out.

As before, lets us start with the simplest scenario. One predictor variable (x): 0= standard drug 1=test drug. Suppose the outcome (y) is change in systolic blood pressure (baseline value – follow up value). We, then, have:

$E(y)=\beta_0+\beta_1 x$

For x=0, we have $E(y)=\beta_0$ and for x=1, we have $E(y)=\beta_0+\beta_1$ and so:

β_1 is the expected change in sbp for those receiving the test drug minus the expected change in sbp for those receiving the standard drug.

Now let us suppose that gender is thought to be a potential confounder or modifier (g=0 (female) g=1 (male)). We can, then, consider:

$E(y)=\beta_0+\beta_1 x+\beta_2 g+\beta_3 gx$

so that, for example,

β_1 is, for the females, the expected change in sbp for those receiving the test drug minus the expected change in sbp for those receiving the standard drug. Sound familiar? All our development of models from the past classes carries forward with appropriate changes in the description of the

characteristic of the outcome.

If $E(y)=\sum_{j=0}^{k}\beta_j x_j$, then we consider possible fits: $Y=\sum_{j=0}^{k} b_j x_j$ and assess the fit by computing the residual: $e_i=y_i-Y_i$ for the ith observation. Then, we adopt the principle of least squares seeking to minimize the residual sum of squares $\sum_{i=1}^{n} e_i^2$. The least squares solution gives us 'fitted values' Y_i for every observation and the solution gives us unbiased estimates b_j of the population characteristics β_j .

Now, suppose we assume that all the y_i are statistically independent and that the [conditional] variance of the observations is constant $Var(y_i)=\sigma^2$ (i.e. the conditional variance does not depend on the conditions [the explanatory variables] and, implicitly, does not depend on $E(y_i)$, then the least squares estimate is the 'Best Linear Unbiased' estimate. Exactly what this BLUE term means we will skip over. Sounds good though, doesn't it?

So far we have discussed three of the assumptions about the conditional distribution. The independence assumption. The mean of the conditional distribution [the conditional mean aka the regression] must a linear function of the explanatory variables and the variance of the conditional distribution [the conditional variance aka the variance about regression] must be constant. Add to these three assumptions our opening argument that the conditional distribution should be symmetrical [to provide meaning and context for the conditional mean] and you have the key ingredients for a standard application of linear regression.

Notice that the homogeneity of conditional variances and the symmetry of the conditional distributions were not required assumptions for any of the binomial regressions or their extensions.

Testing and confidence intervals proceed in ways similar to our earlier work. With linear regression, the Wald tests are replaced with t tests and the likelihood ratio tests are now replaced with F tests. Two sided t tests now give identical p-values to the corresponding one sided F tests.

If we now add one more assumption: that the conditional distributions are normal distributions, then the p-values and confidence intervals are exact. If the conditional distributions are not exactly normally distributed, then the p-values and confidence intervals are approximate as before.

It is worth noting here that these approximations can be valid with large sample sizes even if the conditional distributions are neither normal or symmetrical. Here one is applying the infamous [notorious] Central Limit Theorem. Such p-values and/or confidence limits may well adequately serve the study of the conditional means but one needs to ask if such conditional means are worth studying if the relevant conditional distributions are skewed.

Error
The model discussed above with all the assumptions can be written in an alternate way that can serve the understanding of the issues and challenges of the application of linear regression analysis.

$$y=\sum_{j=0}^{k}\beta_j x_j+\sigma z$$

This equation covers a lot of territory: The right hand side can be written as $\sum_{j=0}^{k} \beta_j x_j$ plus the independent standard normal error z scaled by the constant σ .

Sometimes, you will also see the model written as:

$$y = \sum_{j=0}^{k} \beta_j x_j + \epsilon$$

The only difference between this alternate and the first one is that we now note the errors ϵ are independent normal with mean zero and variance σ^2 . These 2 formulations are identical for us.

It can be helpful to note that the fit can be expressed in multiple ways as well.

$$Y = \sum_{j=0}^{k} b_j x_j$$

or $y = \sum_{j=0}^{k} b_j x_j + e$. The [raw] residual is e. [This is just a rearrangement of e = y – Y] The e's are the fit analogue of the ϵ 's.

or $y = \sum_{j=0}^{k} b_j x_j + sr$. A [standardized] residual is r. The standardized residuals (r) are usually scaled so that they have variance of 1. That way, if the model assumptions hold up, then the standardized residuals should look like standard normal values with mean 0 and variance 1. The r's are the fit analogue of the z's. We see that s is the estimate of σ The standardized residuals are slightly correlated but this is not usually an issue of concern in so far as their graphical assessment is not troubled by this lack of independence . [This can be confusing. The errors ϵ are assumed to be independent but the residuals e are not [quite] independent]

"No aphorism is more frequently repeated in connection with field trials, than that we must ask Nature few questions or, ideally, one question, at a time. The writer is convinced that this view is wholly mistaken. Nature, he suggests, will best respond to a logical and carefully thought out questionnaire; indeed, if we ask her a single question, she will often refuse to answer until some other topic has been discussed." (R.A. Fisher)

The Analysis of Variance

Back in the 1920's, RA Fisher devised a method to test a hypothesis comparing [2 or more] conditional means by means of an ingenious process that involved the comparison of 2 estimates of [the assumed constant] variance. He constructed a test based on the difference between the logarithms of the 2 estimates of variance. Other authors [Snedecor?] suggested that one need not bother with the logarithms and that the test could be based on the ratio of the 2 estimates of variance. And so the F ratio was born. F is for Fisher. Such methods based on the comparison of estimates of variance came to be called the analysis of variance.

The analysis of variance provided considerable clarity in understanding the mechanisms behind a typically complex hierarchy of testing and, for certain [so-called orthogonal] experiments, the calculations could all be done with relative ease compared with the [at the time] almost impossible calculations required for a regression analysis. Unfortunately, the orthogonality [or sometimes called balance] typically required that, for every combination of factors under study, one needed to have

exactly the same number of subjects. This meant designing exact balance and then hoping that there would be no withdrawals or lost-to-follow-ups. Any loss of orthogonality meant tedious [and sometimes fruitless] additional calculations.

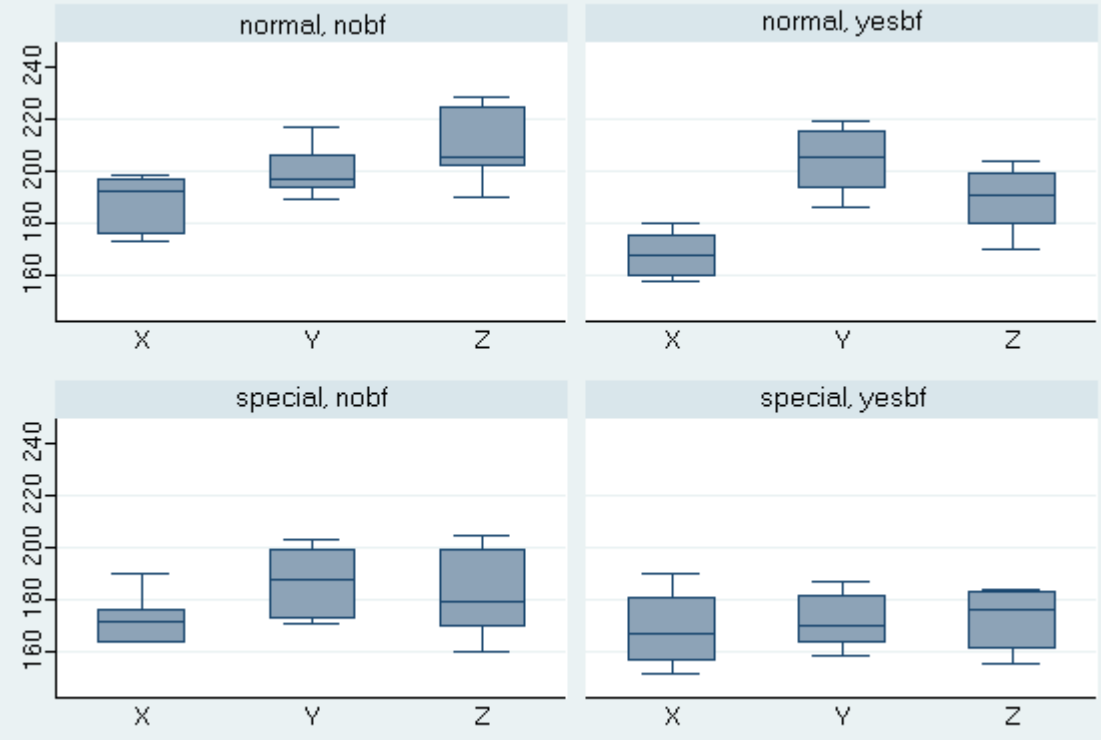
We will now consider a reasonably elaborate application of the analysis of variance and then show that a regression analysis can done, in this case, that reproduces all the key elements of this analysis of variance. We will see a direct relationship here with our earlier work in the assessment of potential confounders/modifiers that are each assumed to take on a finite set of levels.

Consider now the randomized clinical trial starting on page 77 of Rabe-Hesketh & Everitt on treating hypertension. Please read their description of the study, the data management, the data reshaping, the description and their analyses. The data is in bp.dta

```
. anova bp drug diet biofeed diet*drug diet*biofeed drug*biofeed diet*drug*biofeed

                          Number of obs =      72     R-squared     =  0.5840
                          Root MSE      = 12.5167     Adj R-squared =  0.5077

                 Source |  Partial SS    df       MS           F     Prob > F
        ----------------+----------------------------------------------------
                  Model |       13194    11  1199.45455       7.66     0.0000
                        |
                   drug |        3675     2      1837.5      11.73     0.0001
                   diet |        5202     1        5202      33.20     0.0000
                biofeed |        2048     1        2048      13.07     0.0006
               diet*drug |         903     2       451.5       2.88     0.0638
            diet*biofeed |          32     1          32       0.20     0.6529
            drug*biofeed |         259     2       129.5       0.83     0.4425
        diet*drug*biofeed |        1075     2       537.5       3.43     0.0388
                        |
                Residual |        9400    60  156.666667
        ----------------+----------------------------------------------------
                   Total |       22594    71  318.225352
```

Lets use slightly different labels for the values of the variables:
Drug: 1=X 2=Y 3=Z
Diet: 0=normal 1=special
Biofeedback: 0=no biofeedback (nobf) 1=biofeedback (yesbf)
Perhaps it is best, here, to view drug, diet and biofeedback as 'exposures'. Alternatively, one might view drug as the exposure and diet and biofeedback as potential confounders/modifiers.

```
. lab list
dl:
           0 normal
           1 special
bl:
           0 nobf
           1 yesbf
drl:
           1 X
           2 Y
           3 Z

. gen Y=(drug==2)
. gen Z=(drug==3)
. gen S=diet
. gen B=biofeed
. gen YS=Y*S
. gen ZS=Z*S
. gen YB=Y*B
. gen ZB=Z*B
. gen SB=S*B
. gen YSB=YS*B
. gen ZSB=ZS*B
```


normal, nobf
normal, yesbf
special, nobf
special, yesbf
160 180 200 220 240
X
Y
Z
Graphs by diet and biofeed

```
. table drug diet biofeed, c(mean bp)

-----------------------------------------------
          |          biofeed and diet
          | ----- nobf -----    ----- yesbf ----
     drug |  normal  special     normal  special
----------+------------------------------------
        X |     188      173        168      169
        Y |     200      187        204      172
        Z |     209      182        189      173
-----------------------------------------------

. regress bp Y Z S B YS ZS YB ZB SB YSB ZSB

      Source |       SS       df       MS              Number of obs =      72
-------------+------------------------------           F( 11,    60) =    7.66
       Model |       13194    11  1199.45455           Prob > F      =  0.0000
    Residual |        9400    60  156.666667           R-squared     =  0.5840
-------------+------------------------------           Adj R-squared =  0.5077
       Total |       22594    71  318.225352           Root MSE      =  12.517

------------------------------------------------------------------------------
          bp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           Y |         12   7.226494     1.66   0.102    -2.455141    26.45514
           Z |         21   7.226494     2.91   0.005     6.544859    35.45514
           S |        -15   7.226494    -2.08   0.042    -29.45514   -.5448589
           B |        -20   7.226494    -2.77   0.007    -34.45514   -5.544859
          YS |          2   10.21981     0.20   0.846    -18.44266    22.44266
          ZS |        -12   10.21981    -1.17   0.245    -32.44266    8.442657
          YB |         24   10.21981     2.35   0.022     3.557343    44.44266
          ZB |   9.38e-14   10.21981     0.00   1.000    -20.44266    20.44266
          SB |         16   10.21981     1.57   0.123    -4.442657    36.44266
         YSB |        -35   14.45299    -2.42   0.018    -63.91028   -6.089718
         ZSB |         -5   14.45299    -0.35   0.731    -33.91028    23.91028
       _cons |        188   5.109903    36.79   0.000     177.7787    198.2213
------------------------------------------------------------------------------

. test YSB=ZSB=0

 ( 1)  YSB - ZSB = 0
 ( 2)  YSB = 0

       F(  2,    60) =    3.43
            Prob > F =    0.0388
```

$$E(y)=\beta_0+\beta_1 Y+\beta_2 Z+\beta_3 S+\beta_4 B+\beta_5 YS+\beta_6 ZS+\beta_7 YB+\beta_8 ZB+\beta_9 SB+\beta_{10} YSB+\beta_{11} ZSB$$

The regression coefficients are simply differences among means:

β_1 : for those receiving normal diet and no biofeed , mean bp for those receiving drug Y minus mean bp for those receiving drug X [estimated by 200-188 = 12]

β_2 : for those receiving normal diet and no biofeed , mean bp for those receiving drug Z minus mean bp for those receiving drug X [estimated by 209-188 = 21]

β_3 : for those receiving X and no biofeed , mean bp for those receiving special diet minus mean bp for those receiving normal diet [estimated by 173-188 = -15]

β_4 : for those receiving X and normal diet, mean bp for those receiving biofeed minus mean bp for those receiving no biofeed. [estimated by 168-188 = -20]

β_5 : for those with no biofeed , mean bp for those receiving drug Y minus mean bp for those receiving drug X for those receiving special diet minus mean bp for those receiving drug Y minus mean bp for those receiving drug X for those receiving normal diet [estimated by (187-173) – (200-188) = 2]

β_{10} : ...a very long sentence :-) [estimated by ((172-169) – (204-168)) - ((187-173) – (200-188))= -35]

...and so on. It is instructive to complete these interpretations and the estimates. The estimates are based

on differences among averages here in part because there are the same number of participants [6] in each of the 12 groups. Compare this situation with logistic regression and the model with 8 terms. Now we have estimates of means [or estimates of expectations] rather than estimates of log odds.

It is also worthy to notice that the analysis of variance [as detailed above] provides an incomplete break down of the degrees of freedom in that there are several rows in the table with more than one degree of freedom. The regression analysis provides a complete break down with the added decision to view those receiving drug X as the comparison group.

Notice that the F tests from the analysis of variance are not necessarily the same as the corresponding t tests from the regression analysis. The null hypotheses can be different. In this example, the analysis of variance test corresponding to 'diet' has an F=33.20. This test is displayed to consider the null hypothesis that the 'main effect' [sic] of diet [i.e. comparing the diets ignoring drug and biofeedback. Irrelevant here because we are statistically decent :-)] is zero. The arithmetic looks like:

$$t=\frac{\bar{y}_{special}-\bar{y}_{normal}}{s\sqrt{(1/36+1/36)}}=\frac{193-176}{12.517\sqrt{(1/18)}}=5.762 \qquad t^2=(5.762)^2=33.20$$

(Notice, here, that $s=\sqrt{(\text{Mean Square Residual})}$ from the regression analysis)

While the corresponding t test from the regression analysis has a t = -2.08. This test is displayed to consider the null hypothesis that a simple comparison of diets [i.e. for those receiving drug X and no biofeedback.] is zero. Notice that the test ignoring drug and biofeedback has no meaning here but the 'simple' test [specific to drug X and no biofeedback] does mean something.

If we had used the two degree of freedom F test for the three factor interaction, then the p-value is 0.0388. [This test is same whether determined from the analysis of variance or from the regression analysis. See the test command after the regress command above] There is evidence [at the 5% level] that the drug-mean blood pressure associations depend on the combinations of diet and biofeedback. [i.e there are complex interactions that require detailed study] If we were to view the drug comparisons as primary, then we might say that diet and biofeedback modify the drug - mean blood pressure association .

If we had decided apriori to use the t tests from the regression analysis, then we could note that since the p-value [for YSB] = 0.0118, there is evidence that the comparison between Y and X depends on the combinations of diet and biofeedback. If we proceed to drop ZSB and then assess this model, we get no evidence for SB but modest evidence for ZS [Try this out!]. We might then note that the comparison between Z and X depends on diet (but not on biofeedback). All of this is getting complicated. Time to reconsider the boxplots above given the testing carried out.

You will have noticed that Stata's command for the analysis of variance is anova. The abbreviation ANOVA [ANalysis Of VAriance] was apparently first used by John Tukey. There are many statisticians [me included] who do not like to use these abbreviations. ANOVA [spoken A No Va] seems to have crept into the English language now. Sigh.....

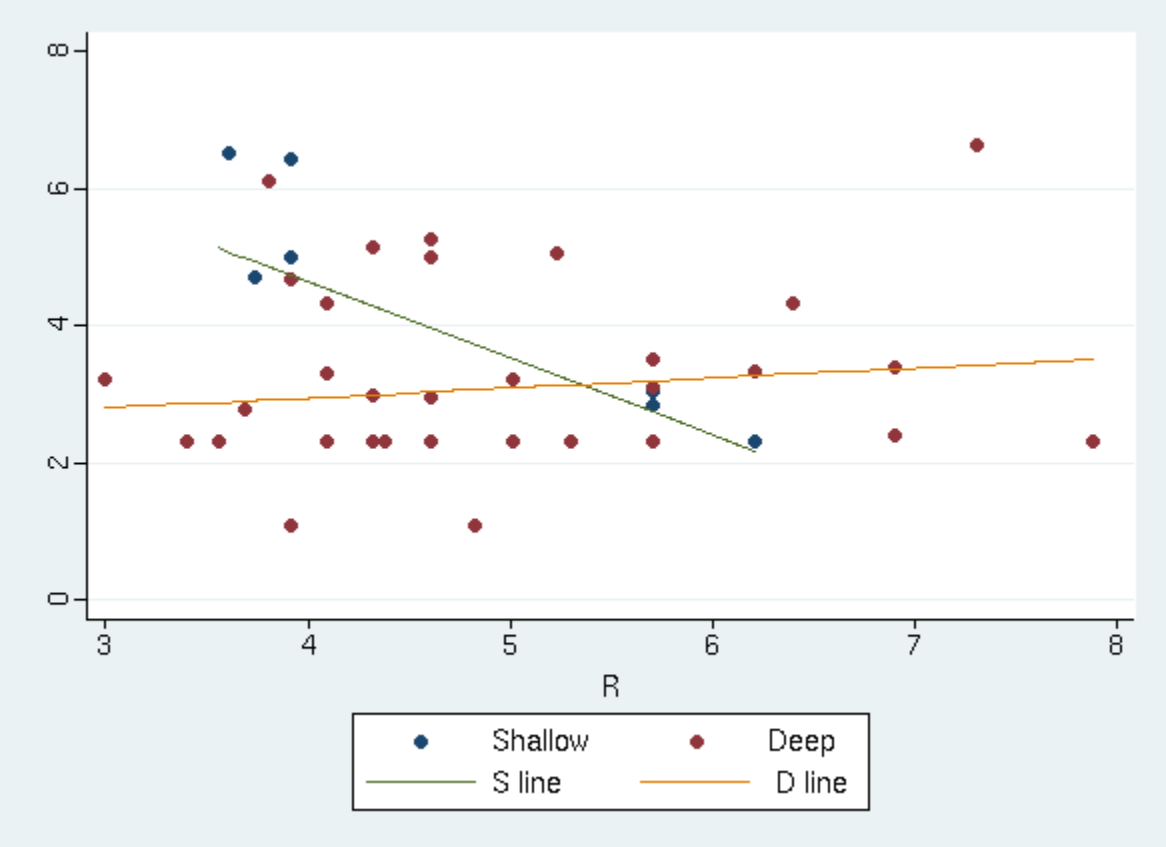
Lets now consider wells.dta from Hamilton(1992) pages 86 to 92. The data is in wells.dta Hamilton argues that both chlorine and distance from the road should analyzed on the logarithmic scale. We will not debate or discuss this matter here. We will take a future class for a discussion of the issue of data transformations. We will transform both variables as in Hamilton. We will construct an indicator W for the depth of the well [0=shallow 1=deep]. We will describe the log of the distance from the road as the exposure and the depth of the well as a potential confounder/modifier.

```
use wells.dta
lab drop deeplbl
lab def dl 0 "S" 1 "D"
rename deep well
lab val well dl
rename droad road
gen lc=log(chlor)
gen W=well
gen R=log(road)
gen WR=W*R
regr lc W R WR
```

```
      Source |       SS       df       MS              Number of obs =      52
-------------+------------------------------           F(  3,    48) =    3.81
       Model |  18.4831272     3   6.1610424           Prob > F      =  0.0157
    Residual |  77.5390714    48  1.61539732           R-squared     =  0.1925
-------------+------------------------------           Adj R-squared =  0.1420
       Total |  96.0221986    51  1.88278821           Root MSE      =   1.271

------------------------------------------------------------------------------
          lc |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           W |  -6.717366   2.094713    -3.21   0.002    -10.92907   -2.505663
           R |  -1.109424   .3844204    -2.89   0.006    -1.882354   -.3364954
          WR |   1.255847   .4268777     2.94   0.005     .3975521    2.114143
       _cons |   9.073459   1.879384     4.83   0.000     5.294704    12.85221
------------------------------------------------------------------------------
```

```
predict lch
sort W R
twoway (scatter lc R if W==0) (scatter lc R if W==1) (line lch R if W==0) (line lch R if
W==1),legend(order (1 "Shallow" 2 "Deep" 3 "S line" 4 " D line"))
```

Here, we can see that the rate of change of the mean of the log of the chlorine per unit change in the log of the distance of well from the road depends on the depth of the well. (p=0.005) Given this detection of the depth of the well as a modifier, we would argue that a regression analysis that excludes WR would be inappropriate and irrelevant. Hamilton does show the other fits for illustrative purposes but, in the end, he notes:



"Among deep bedrock wells, we see virtually no relationship between distance and chlorine. [snip] The closer a shallow well is to the road, the higher its chlorine concentration tends to be." - Hamilton p 90
He also offers a plausible explanation for this finding.

The same fit can be determined using S as the indicator for shallow wells. This regression analysis explicitly displays the estimated rate of change of the mean log(chlorine) per unit change of log(distance)[0.1464] and the p-value of 0.434 for the deep wells. While the first analysis gave us the estimate for the shallow wells [-1.1094] with p-value of 0.006

```
. gen S = 1-W

. gen SR = S*R

. regr lc S R SR
```

Source	SS	df	MS		
Model	18.4831272	3	6.1610424	Number of obs	= 52
Residual	77.5390714	48	1.61539732	F(3, 48)	= 3.81
				Prob > F	= 0.0157
				R-squared	= 0.1925
				Adj R-squared	= 0.1420
Total	96.0221986	51	1.88278821	Root MSE	= 1.271

lc	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
S	6.717366	2.094713	3.21	0.002	2.505663	10.92907
R	.1464229	.1855951	0.79	0.434	-.2267411	.5195869
SR	-1.255847	.4268777	-2.94	0.005	-2.114143	-.3975521
_cons	2.356093	.9250614	2.55	0.014	.496132	4.216053

A regression analysis that begins with the assessment of a variable as a modifier and then possibly proceeding to an assessment of this variable as a confounder is very familiar to us from our work with stratified analyses.

Other Characteristics of the Conditional Distributions
It is worth emphasizing the 'Linear' Regression is concerned with the study of conditional means. We are assuming that the conditional variances do not depend on the condition(s) and [rather implicitly] assuming that all of our interest is with the mean of these distributions.

One can [in principle] study any percentile of these distributions. The Stata command 'qreg' allows the investigator to specify any percentile. The most commonly considered percentile is the median [quantile = 0.5]. One now sees the explicit study of the quartiles [quantile = 0.25 or 0.75] and, in some health research contexts [notably health sociology], other percentiles are studied to considerable advantage.

We can have models like:
$Q_q(y)=\sum \beta_{qi} x_i$ where Q_q is the qth quantile of the conditional distribution
Typically, the regression coefficients will depend on q.

As a very brief example, consider:

$Q_{0.5}(y)=\beta_0+\beta_1 W+\beta_2 R+\beta_3 WR$

```
. qreg lc W R WR,quantile(0.5)

Median regression                                       Number of obs =         52
  Raw sum of deviations 51.74384 (about 2.3025851)
  Min sum of deviations 44.52187                        Pseudo R2     =     0.1396

------------------------------------------------------------------------------
          lc |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           W |  -7.563186   2.672955    -2.83   0.007    -12.93752    -2.18885
           R |  -1.215242   .4887267    -2.49   0.016    -2.197893   -.2325915
          WR |   1.243673   .5440056     2.29   0.027     .1498762    2.337469
       _cons |   9.764691   2.389528     4.09   0.000     4.960222    14.56916
------------------------------------------------------------------------------
```

Now, for this example, the interpretation is in terms of conditional medians rather than conditional means. So, for the shallow wells, the estimate of the rate of change of the median log(concentration) per unit change in log (distance) is -1.2152
The tests are Wald tests and likelihood ratio tests. One is assuming that the conditional distributions have constant scale and shape. All that changes [as a function of the explanatory variables] is the location.
The help file on qreg is 'helpful'. There are several books entirely devoted to quantile regression. The book by Koenker (2005) and chapter 7 in Cameron & Trivedi (2009) give the theory and examples.