Biost 536: Categorical Data Analysis in Epidemiology
Emerson, Fall 2014

Homework #1 Key
September 26, 2014

Written problems: To be submitted as a MS-Word compatible file to the class Catalyst dropbox by 1:30 pm on Thursday, September 26, 2014. See the instructions for peer grading of the homework that are posted on the web pages.

> *On this (as all homeworks) Stata / R code and unedited Stata / R output is **TOTALLY** unacceptable. Instead, prepare a table of statistics gleaned from the Stata output. The table should be appropriate for inclusion in a scientific report, with all statistics rounded to a reasonable number of significant digits. (I am interested in how statistics are used to answer the scientific question.)*
>
> ***In all problems requesting "statistical analyses" (either descriptive or inferential), you should present both***
>
> - ***Methods:** A brief sentence or paragraph describing the statistical methods you used. This should be using wording suitable for a scientific journal, though it might be a little more detailed. A reader should be able to reproduce your analysis. DO NOT PROVIDE Stata OR R CODE.*
> - ***Inference: A paragraph providing full statistical inference in answer to the question. Please see the supplementary document relating to "Reporting Associations" for details.***
>
> *Keys to past homeworks from quarters that I taught Biost 517 (e.g. HW #8 from 2012) or Biost 518 (e.g., HW #1 from 2014) or Biost 536 (e.g. HW #3 from 2013) might be consulted for the presentation of inferential results. Note that the requirement to provide a paragraph describing your statistical methods was new this year in Biost 518, and other past keys do not give explicit examples of a separate paragraph. However, many past keys provide this information as an introductory sentence.*

All questions of treatment benefit relate to any benefit of idarubicin to more frequently induce complete remission in patients with acute myelogenous leukemia (AML) relative to daunorubicin. The data can be found on the class web page (follow the link to Datasets) in the file labeled leukemia.txt. Documentation is in the file leukemia.doc. The data is in free-field format. The file initleuk.txt contains Stata and R code that might be useful in reading in the file.

In this homework, we are ultimately interested in inference for the primary endpoint (induction of complete remission) by treatment and sex. We will ignore the sequential aspect of the clinical trial and just analyze the data as if we had always planned on getting data on 130 subjects. We will use an intent to treat approach that analyzes all subjects according to how they were randomized.

Instructions for grading: On this key I place the total points to be awarded for a particular problem just prior to the answer for the problem. I also sometimes provide very specific criteria for awarding points. Remember the purpose of peer grading is to have the grader focus more closely on answers to the problems that were potentially different from those that he/she provided, and to identify areas where the answers on the paper being graded might not be correct. It is not of value to anyone to be overly "lenient" in the grading. So while it is not appropriate to capriciously deduct points, it is equally not appropriate to give points only because "Well, they

tried" (unless, of course, such a criterion is specified in the grading instructions). We welcome students' questions regarding the appropriateness of answers during the grading process. And grades assigned by the peer grader can always be appealed.

In providing answers below, the required answers are in* bold face type*. Further verbage that is in *regular italics type* ***is additional information that you are in fact responsible for knowing on exams, but was not really required for the homework assignment. That additional information may help you decide whether the answers you are grading are valid, however.***

In each problem requiring both description of methods and results, full credit should only be given when both aspects are appropriately addressed. The grader should be able to determine the exact analysis method from their description.

At the end of this document, I provide Stata and R code that I used to produce the answers. This is for your later reference when trying to understand Stata or R. I did not expect (and in fact do not tolerate) the Stata and R output as solutions to the homework assignment. It should not have been included in a homework set turned in by a student.

Ans: *The primary outcome for this clinical trial was induction of complete remission within the first two courses of chemotherapy: a binary outcome. Important secondary outcomes were the induction of complete remission with the course of chemotherapy and overall survival.*

- *Overall survival is of course the most important clinical endpoint, but the investigators likely worried that there would not be sufficient power with the planned sample size.*
- *Obviously, it is preferable to induce complete remission in the first course of chemotherapy. However, if that does not obtain, clinicians were willing to try a second course. And if there was an advantage of one drug after the second course of chemotherapy, that drug would be considered better, no matter what its ability to induce CR after the first remission. In such a setting, we would not want to look at time to remission, lest that analysis choose the wrong drug.*

.

1. Provide suitable descriptive statistics for this dataset as might be presented in Table 1 of a manuscript appearing in the medical literature.

<u>Instructions for grading</u>: This problem is worth 10 points. The table should present descriptive statistics in an organized and well-formatted fashion. Key points to consider in your grading:

- ***Columns should be by treatment group, and sample size should be apparent. It is not necessary to present a "Combined Group" column, but it is not wrong to do so (I just think it is a waste of space, if it leads to messier tables within the columns).***
- ***P values are of no use in the table and should not be included for an RCT (yes, I know many papers will have them, but lots of people run red lights, too).***
- ***Given our motivation for presenting descriptive statistics ("Materials and Methods" and to assess possible randomization imbalances that lead to what I call "conditional confounding"), the best descriptive statistics to present for continuous data will be either mean or geometric mean (depending upon whether adjustment would be on the log scale), standard deviation (if mean is presented), minimum and maximum. Medians are of little use in describing importance imbalances in randomization (unless they are very good surrogates for the geometric mean), and unless you have enough space to present both medians and means, it is generally best to suppress the medians. In my table below, I do present medians figuring that for WBC it is a pretty good surrogate for the geometric mean, but personally I think the GM would be better. (And again, I invoke the problem with red lights re the prevalence of medians***

in reports of RCTs in the medical literature. For a full discussion of the merits of various descriptive statistics vis a vis the goal of presenting descriptive statistics, see lectures from Biost 517, as well as discussion of confounding from Biost 518. The latter will be covered again in great detail this quarter.)

- ***The only variables that should be included in the table are those variables measured pre-randomization: sex, age, FAB, Karnofsky score, WBC (geometric means are preferable for this variable in leukemia, though medians probably are reasonable surrogates for the geometric mean), platelets, and hemoglobin.***
- ***FAB is a nominal variable, and it is erroneous to present means, SD, min, max for this variable. Sex is of course binary. All other variables are probably best described with mean, sd, etc. If there are important thresholds (e.g., Karnofsky $\geq$ 80, WBC $\leq$ 2 or WBC $\geq$ 25, platelets $\leq$ 30, hemoglobin $\leq$ 7), it is okay to also present the proportion above or below such thresholds (there are none that I would have noted, though the examples I gave were ones that I might have considered).***
- ***There should be some way that missing data was noted in the table. This could be in footnotes or as a separate statistic in the table. (Given the pattern of missing data, I find it easiest to do it as footnotes.)***
- ***It would be unusual, but not necessarily wrong, to present descriptive statistics for the time of randomization. Hopefully, randomization was blocked to guarantee balance across treatment arm over time and thus avoid imbalance in any patient characteristics that might have varied over the time of randomization.***

Ans: *Methods:* Descriptive statistics are provided by treatment arm for variables measured pre-randomization. Continuous variables are summarized by means, standard deviations, medians, minima, and maxima. Binary and categorical data are summarized by their frequency distribution.

***Results:* Table 1 contains descriptive statistics for baseline variables for the 130 randomized patients (65 received daunorubicin (DNR), 65 received idarubicin (IDR)). One patient on the DNR arm was missing data for the hematologic variables, and the French-American-British classification of cell type was missing for 8 DNR and 4 IDR patients. Patients ranged in age from 17 to 61 and were equally likely to be male or female.**

		Daunorubicin (DNR) (n = 65)	Idarubicin (IDR) (n = 65)
Males (n (%))		35 (53.8%)	30 (46.2%)
Age (years)		39.8 (sd 13.4; mdn 40.0; range 19-60)	38.0 (sd 12.5; mdn 36.0; range 17 - 61)
Perf status (Karnofsky)		79.5 (sd 12.6; mdn 80.0; range 40-100)	79.5 (sd 11.6; mdn 80.0; range 30-100)
White blood cells (10^3/mm^3)		43.3 (sd 55.0; mdn 16.7; range 0.7-215)	29.0 (sd 36.3; mdn 11.8; range 2.8-13.7)
Platelets (10^3/mm^3)		93.6 (sd 92.4; mdn 62.0; range 11-457)	66.6 (sd 57.8; mdn 50.0; range 11-370)
Hemoglobin (g/dl)		9.64 (sd 1.49; mdn 9.45; range 6.4-13.9)	9.22 (sd 1.82; mdn 9.2; range 2.8-13.7)
Fr-Am-Br class (n (%))	**0**	1 (1.8%)	0 (0.0%)
	1	6 (10.5%)	13 (21.3%)
	2	15 (26.3%)	15 (24.6%)
	3	9 (15.8%)	11 (18.0%)
	4	12 (21.1%)	8 (13.1%)
	5	13 (22.8%)	12 (19.7%)
	6	1 (1.8%)	2 (3.3%)

2. Perform an analysis to assess whether subjects taking idarubicin have better primary clinical outcomes than patients on daunorubicin.

***Instructions for grading*: This problem is worth 10 points. Key points to consider in your grading:**

- ***The answer should include point estimates of the proportion experiencing a complete remission rate in each arm, as well as a point estimate of association (RD, RR, or OR), a confidence interval for that measure of association, and a p value. Note that interpretation of the clinical importance of an RR or OR is impossible without knowing the probability of remission on each arm.***
- ***Method of analysis should ideally be based on an analysis of the variable "cr" using either RD, RR, or OR. In general, the methods of estimation (point estimate and CI) use "maximum likelihood methods". P values might be based on***
    - ***an unconditional exact test (ideal and easy if you use R),***
    - ***the chi square test (generally next best), or***
    - ***Fisher's exact test (which is slightly conservative in these sample sizes).***
- ***Methods based on the t test that presumes equal variances (for p values) or that allows for the possibility of unequal variances (for CI) are also okay. And because this inference is valid, so is inference based on linear regression. If you use linear regression, the p value based on robust standard errors is valid in large sample sizes, but the p value based on classical linear regression is generally more efficient under the null hypothesis. For CI, the use of the robust standard errors is probably more efficient, though it would not make much difference when the sample sizes are equal.*** *(For a two group comparison, I would generally just describe the use of the chi square statistic and Wald based CI, because that will raise fewer questions. But when I want to answer problems 4 or 7 using the RD, I will use linear regression. In any case, you may have to defend the appropriateness of these methods based on the central limit theorem and pointing out that distinctions between the chi square test and the t test are quite negligible.)*
- ***As noted above, there are good clinical reasons to prefer an analysis based on the induction of remission within two courses of chemotherapy as the primary endpoint. However, it is not entirely wrong to analyze time to remission as measured by the variable crchemo. To perform such an analysis, you could perform a proportional hazards analysis by regarding all subjects with NA for crchemo as censored. While you could consider the date of remission, that is probably less clinically important, because it is the number of courses of chemotherapy that matter the most. I do not include an analysis based on this approach in this key. If you want to give credit to someone doing such an analysis, make certain that they interpreted the variables correctly and that they censored people who did not achieve a remission (or put their time to remission at some number larger than 2).***
- ***It is not appropriate to analyze the data based on survival time: the question specifically asked for analyses related to induction of complete remission. Hence, any analysis based on the variables fudate and status are not correct.***
- ***An analysis based on the variable "crchemo" is less desirable, but would be acceptable. might be based on RD, RR, or OR.***

Ans: ***Methods:*** *I provide description of methods in approximate order of my personal preference:*

- **Treatment arms were compared using the difference in the probability of inducing a complete remission: IDR arm minus DNR arm. The probability of inducing a complete**

remission was estimated for each treatment arm using the sample proportion, and the difference in those sample proportions was used as the estimated treatment effect. A 95% confidence interval for the difference in remission probabilities was calculated using Wald type CI based on the approximate normal distribution for the maximum likelihood estimates for a binomial distribution. P values testing the null hypothesis of no difference in remission probabilities were computed using *(Then we can describe the method for p values using the relevant approach:*

- **an unconditional exact test based on the chi square statistic (or Fisher's exact test), or**
- **the chi square test, or**
- **Fisher's exact test (without Yates' continuity correction, though you generally do not have to state the last point, because the presumption would be that you would instead have to indicate when you did use it).**

- *If you use the odds ratio, the first sentence describing the methods would be something like:* **Treatment arms were compared using the odds ratio comparing the odds of remission on the IDR arm to the odds of remission on the DNR arm. The probability of inducing a complete remission was estimated for each treatment arm using the sample proportion, and the odds ratio and its 95% confidence interval was calculated using Wald type CI based on the approximate normal distribution for the maximum likelihood estimates for a binomial distribution.** *The rest of the description can read just like the previous description of p values. I note that you could also use CI for the odds ratio based on Fishers exact test.*
- *If you use the rate ratio, just substitute "rate ratio" for "odds ratio" above.*

Results: **Complete remissions were induced in 51 of 65 patients (78.5%) on the IDR arm, and 38 of 65 patients (58.5%) on the DNR arm.** *Now choose the appropriate one of the following statements about the point and interval estimates of the treatment effect:*

- **We thus estimate that IDR is associated with a 20.0% absolute increase in remission probabilities. Based on a 95% CI, such an observation is judged to be consistent with a true absolute increase in remission probabilities between 4.40% and 35.6% higher.** *or*
- **We thus estimate that the odds of remission with IDR use is 2.59-fold higher than it is with DNR. Based on a 95% CI, such an observation is judged to be consistent with a true 1.21-fold to 5.55-fold relative increase in the odds of remission.** *or*
- **We thus estimate that the probability of remission with IDR use is 1.34-fold higher than it is with DNR. Based on a 95% CI, such an observation is judged to be consistent with a true 1.05-fold to 1.71-fold relative increase in the probability of remission.**

Based on this data, we reject the null hypothesis of equal likelihood of remission between treatments (two-sided p value 0.0175 from exact unconditional test *or 0.0141 from chi square test or 0.023 from Fisher's exact test***).**

3. Is the analysis of treatment effect confounded by sex? Provide your reasoning.

Instructions for grading: This problem is worth 10 points. Tests of statistical significance have no place in judging confounding, so arguments based on considering whether there is a statistically significant association between treatment arm and sex or between induction of complete remission and sex are wrong.

Ans: The data arise from a randomized clinical trial, and the above analyses were performed according to randomization. Hence we are assured of no confounding "on average". That is, the estimated treatment effect is consistent for the true treatment effect (and unbiased if we use RD). We can be concerned about "conditional confounding" that would result from chance imbalances in this particular study. However, the treatment arms look relatively balanced with respect to sex.

4. Perform an analysis to assess any treatment benefit of idarubicin over daunorubicin adjusted for sex.

***Instructions for grading**: This problem is worth 10 points. Esthetically, I tend to prefer to use the same measure of association for adjusted and unadjusted analyses. Hence, I think it best (but not crucial) that the analysis model here agree with the analysis model used in problem 2. In any case, the description of the methods and results will tend to be very similar.*

Ans: ***Methods:*** *Depending on the method of analysis:*

- **RD: Treatment arms were compared using the difference in the probability of inducing a complete remission: IDR arm minus DNR arm. Crude (unadjusted) estimates of the probability of inducing a complete remission were estimated for each treatment arm using the sample proportions, Point and interval estimates for the difference in remission probabilities were based on a linear regression of the binary indicator of complete remission on a model that included an indicator of treatment arm and an indicator of sex. The sex adjusted estimate of treatment effect was based on the regression parameter for the indicator of treatment. A 95% confidence interval for the sex adjusted difference in remission probabilities was calculated using Wald type inference based on the Huber-White sandwich estimator of standard errors in order to allow for the mean-variance relationship associated with binary random variables. Confidence intervals and p values assumed the approximate normal distribution for the regression parameter estimates.**
    - *An alternative approach could have been based on stratified RD analysis using internal standardization. See the Stata output.in the Appendix to this key.*
- **OR: Treatment arms were compared using the odds ratio comparing the odds of IDR inducing a complete remission compared to DNR. Crude (unadjusted) estimates of the probability of inducing a complete remission were estimated for each treatment arm using the sample proportions, Point and interval estimates for the treatment odds ratio were based on a logistic regression of the binary indicator of complete remission on a model that included an indicator of treatment arm and an indicator of sex. The sex adjusted estimate of treatment effect was based on the regression parameter for the indicator of treatment. A 95% confidence interval and p value were computed assuming the approximate normal distribution for the regression parameter estimates.**
    - *An alternative approach could have been based on the Mantel-Haenszel statistic. See the Stata output in the Appendix to this key.*
- **RR: Treatment arms were compared using the risk ratio comparing the probability of IDR inducing a complete remission compared to DNR. Crude (unadjusted) estimates of the probability of inducing a complete remission were estimated for each treatment arm using the sample proportions, Point and interval estimates for the ratio of remission probabilities were based on a Poisson regression of the binary indicator of complete remission on a model that included an indicator of treatment arm and an indicator of sex. The sex adjusted estimate of treatment effect was based on the regression parameter for the indicator of treatment. A 95% confidence interval and p value were computed assuming the approximate normal distribution for the regression parameter estimates along with the**

standard errors based on Huber-White sandwich estimator in order to allow for the mean-variance relationship of binary response variables.

- *An alternative approach could have been based on the Mantel-Haenszel statistic. See the Stata output in the Appendix to this key.*

Results: **Complete remissions were induced in 51 of 65 patients (78.5%) on the IDR arm, and 38 of 65 patients (58.5%) on the DNR arm.** *Now choose the appropriate one of the following statements about the point and interval estimates of the treatment effect:*

- **RD: Based on a linear regression model adjusting for patient sex, IDR use is estimated to be associated with an absolute increase of 18.5% in the probability of an induced complete remission (95% CI 2.85% to 34.3% absolute increase). Based on a two-sided p value of 0.021, we reject the null hypothesis that the treatments are equal with respect to induction of CR.**
  - *With an "internally" standardized stratified analysis, the estimate would have been 18.3% with a 95% CI of 3.10% to 33.6%.*
- **OR: Based on a logistic regression model adjusting for patient sex, IDR use is estimated to be associated with 2.51-fold higher odds of an induced complete remission (95% CI 1.14-fold to 5.51-fold higher odds). Based on a two-sided p value of 0.022, we reject the null hypothesis that the treatments are equal with respect to induction of CR.**
  - *With a stratified analysis using the Mantel-Haenszel statistic, the estimate would have been an OR of 2.51 with a 95% CI of 1.14 to 5.51, and a p value of 0.021.This is not surprising, as the Mantel-Haenszel statistic is closely related to the score test from logistic regression, while the prior analysis was based on Wald statistics.*
- **RR: Based on a Poisson regression model adjusting for patient sex, IDR use is estimated to be associated with 1.31-fold higher odds of an induced complete remission (95% CI 1.04-fold to 1.67-fold higher odds). Based on a two-sided p value of 0.024, we reject the null hypothesis that the treatments are equal with respect to induction of CR.**
  - *With a stratified analysis using the Mantel-Haenszel statistic, the estimate would have been an RR of 1.31 with a 95% CI of 1.04 to 1.66.*

5. Perform an analysis to assess whether males taking idarubicin have more frequent complete remission than males taking daunorubicin.

Instructions for grading: This problem is worth 10 points. This problem is best answered by a "disaggregated" analysis which restricts the analysis to males. It could be answered using a model with a sex-treatment interaction, but that model has more assumptions about the variability of the statistics.

Ans: ***Methods:*** **Analysis of the data was restricted to males.** *And now the rest of the methods reads just like problem 2.*

Results: **Complete remissions were induced in 21 of 30 males (70.0%) on the IDR arm, and 17 of 35 males (48,6%) on the DNR arm.** *Now choose the appropriate one of the following statements about the point and interval estimates of the treatment effect:*

- **We thus estimate that IDR is associated with a 21.4% absolute increase in remission probabilities in males. Based on a 95% CI, such an observation is judged to be consistent with a true absolute difference in remission probabilities between 1.88% lower and 44.7% higher.** *or*

- **We thus estimate that the odds of remission with IDR use in males is 2.47-fold higher than it is with DNR. Based on a 95% CI, such an observation is judged to be consistent with a true relative 10.2% lower odds to 6.78-fold higher odds of remission.** *or*
- **We thus estimate that the probability of remission with IDR use in males is 1.44-fold higher than it is with DNR. Based on a 95% CI, such an observation is judged to be consistent with a true relative 4.70% lower to 118% higher probability of remission.**

Based on this data, we cannot reject the null hypothesis of equal likelihood of remission between treatments in males (two-sided p value 0.0929 *(or 0.102)* **from exact unconditional test** *or 0.0805 from chi square test or 0.129 from Fisher's exact test***).**

6. Repeat problem 5 for females.

<u>Instructions for grading</u>: This problem is worth 10 points. This problem is best answered by a "disaggregated" analysis which restricts the analysis to females. It could be answered using a model with a sex-treatment interaction, but that model has more assumptions about the variability of the statistics.

Ans: ***<u>Methods:</u>*** **Analysis of the data was restricted to females.** *And now the rest of the methods reads just like problem 2.*

<u>Results:</u> **Complete remissions were induced in 30 of 35 females (85.7%) on the IDR arm, and 21 of 30 females (70.0%) on the DNR arm.** *Now choose the appropriate one of the following statements about the point and interval estimates of the treatment effect:*

- **We thus estimate that IDR is associated with a 15.7% absolute increase in remission probabilities in females. Based on a 95% CI, such an observation is judged to be consistent with a true absolute difference in remission probabilities between 4.37% lower and 35.8% higher.** *or*
- **We thus estimate that the odds of remission with IDR use in females is 2.57-fold higher than it is with DNR. Based on a 95% CI, such an observation is judged to be consistent with a true relative 21.9% lower odds to 8.40-fold higher odds of remission.** *or*
- **We thus estimate that the probability of remission with IDR use in females is 1.22-fold higher than it is with DNR. Based on a 95% CI, such an observation is judged to be consistent with a true relative 6.57% lower to 60.5% higher probability of remission.**

Based on this data, we cannot reject the null hypothesis of equal likelihood of remission between treatments in males (two-sided p value 0.152 *(or .139)* **from exact unconditional test** *or 0.124 from chi square test or 0.143 from Fisher's exact test***).**

7. Perform an analysis to assess whether any treatment benefit of idarubicin over daunorubicin differs by sex.

<u>Instructions for grading</u>: This problem is worth 10 points. Esthetically, I tend to prefer to use the same measure of association for adjusted, unadjusted and interaction analyses. Hence, I think it best (but not crucial) that the analysis model here agree with the analysis model used in problems 2 and 4. In any case, the description of the methods and results will tend to be very similar to those earlier problems..

Ans: ***<u>Methods:</u>*** *Depending on the method of analysis:*

- **RD: Treatment arms were compared using the difference in the probability of inducing a complete remission: IDR arm minus DNR arm allowing for different effects in each sex. Crude (unadjusted) estimates of the probability of inducing a complete remission were estimated for each sex on each treatment arm using the sample proportions, Point and interval estimates for the difference in remission probabilities for each sex were based on a linear regression of the binary indicator of complete remission on a model that included an indicator of treatment arm, an indicator of sex, and the multiplicative interaction of those two terms. The estimate of differential treatment effect across sexes was based on the regression parameter for the sex-treatment interaction. A 95% confidence interval for effect modification in remission probabilities was calculated using Wald type inference based on the Huber-White sandwich estimator of standard errors in order to allow for the mean-variance relationship associated with binary random variables. Confidence intervals and p values assumed the approximate normal distribution for the regression parameter estimates..**
- **OR: Treatment arms were compared using the odds ratio comparing the odds of IDR inducing a complete remission compared to DNR allowing for different effects in each sex. Crude (unadjusted) estimates of the probability of inducing a complete remission were estimated for each sex on each treatment arm using the sample proportions, Point and interval estimates for the treatment odds ratio for each sex were based on a logistic regression of the binary indicator of complete remission on a model that included an indicator of treatment arm, an indicator of sex, and the multiplicative interaction of those two terms.. The estimate of differential treatment effect across sexes was based on the regression parameter for the sex-treatment interaction. A 95% confidence interval and p value were computed assuming the approximate normal distribution for the regression parameter estimates.**
    - *An alternative approach could have been based on the Mantel-Haenszel test for homogeneity of OR. See the Stata output in the Appendix to this key.*
- **RR: Treatment arms were compared using the risk ratio comparing the probability of IDR inducing a complete remission compared to DNR allowing for different effects in each sex. Crude (unadjusted) estimates of the probability of inducing a complete remission were estimated for each sex on each treatment arm using the sample proportions, Point and interval estimates for the ratio of remission probabilities for each sex were based on a Poisson regression of the binary indicator of complete remission on a model that included an indicator of treatment arm, an indicator of sex, and the multiplicative interaction of those two terms.. The estimate of differential treatment effect across sexes was based on the regression parameter for the sex-treatment interaction. A 95% confidence interval and p value were computed assuming the approximate normal distribution for the regression parameter estimates along with the standard errors based on Huber-White sandwich estimator in order to allow for the mean-variance relationship of binary response variables.**
    - *An alternative approach could have been based on the Mantel-Haenszel statistic test for homogeneity of RR. See the Stata output in the Appendix to this key.*

Results: **Complete remissions were induced in 21 of 30 males (70.0%) on the IDR arm, and 17 of 35 males (48,6%) on the DNR arm. For females, complete remissions were induced in 30 of 35 females (85.7%) on the IDR arm, and 21 of 30 females (70.0%) on the DNR arm.** *Now choose the appropriate one of the following statements about the point and interval estimates of the treatment effect:*

- **RD: We thus estimate that IDR is associated with a 21.4% absolute increase in remission probabilities in males, and associated with a 15.7% absolute increase in remission**

probabilities in females. Based on a linear regression model adjusting for treatment, patient sex, and their multiplicative interaction we estimate a treatment effect in males that is 5.7% higher than that in females (95% CI -25.8% to 37.2%). Based on a two-sided p value of 0.721, we cannot reject the null hypothesis that the treatment effect is similar in males and females.

- **OR: We thus estimate that the odds of remission with IDR use in males is 2.47-fold higher than it is with DNR in males, while in females we estimate that the odds of remission with IDR use is 2.57-fold higher than it is with DNR. Thus the estimate of an OR in males is only 0.961 times that of the OR in females (95% CI for ratio of OR 0.194 to 4.75), and based on a two-sided p value of 0.961, we cannot reject the null hypothesis that the treatment effect is similar in males and females.**
    - *With a stratified analysis using the Mantel-Haenszel test of homogeneity, the p value of 0.961 is similar to that from the logistic regression. This is not surprising, as the Mantel-Haenszel statistic is closely related to the score test from logistic regression, while the prior analysis was based on Wald statistics.*
- **We thus estimate that the probability of remission with IDR use in males is 1.44-fold higher than it is with DNR in males, while in females we estimate that the probability of remission with IDR use is 1.22-fold higher than it is with DNR. Thus the estimate of an RR in males is 1.18 times that of the RR in females (95% CI for ratio of RR 0.458 to 1.05), and based on a two-sided p value of 0.520, we cannot reject the null hypothesis that the treatment effect is similar in males and females.**
    - *With a stratified analysis using the Mantel-Haenszel test of homogeneity, the p value of 0.4511 is similar to that from the poisson regression..*

8. Use the analysis you performed in problem 7 to answer the question of whether idarubicin use is associated with more frequent induction of remission.

***Instructions for grading*: This problem is worth 10 points. Note that I guessed that few if any students would have answered question 7 with the stratified analysis. Had anyone done that, this question was just the stratified estimate of effect. I answer the question with the regression models. Credit should only be given if it you can verify that they performed the "multiple partial" test that considers both relevant parameters simultaneously.**

Ans: *Methods:* In the regression model fit with the indicator of treatment, the indicator of sex, and the multiplicative interaction, we perform a Wald test that both the regression parameter for the treatment effect and the regression parameter for the treatment-sex interaction are simultaneously equal to zero.*:

Results: *Choose the appropriate one of the following statements about "multiple partial" test:*

- **RD: Based on a p value of 0.071, we cannot reject the null hypothesis that the likelihood of complete remission is similar between IDR and DNR after adjusting for sex and the sex-treatment interaction.**
- **OR: Based on a p value of 0.072, we cannot reject the null hypothesis that the likelihood of complete remission is similar between IDR and DNR after adjusting for sex and the sex-treatment interaction.**
- **RR: Based on a p value of 0.078, we cannot reject the null hypothesis that the likelihood of complete remission is similar between IDR and DNR after adjusting for sex and the sex-treatment interaction.**

9. For each of the analysis models used in problems 4 and 7, provide estimates of the probability of inducing a complete remission by all combinations of treatment group and sex. How do these estimates compare to descriptive statistics for those groups?

***Instructions for grading*: This problem is worth 10 points.**

Ans: ***Methods:*** *There was no need to supply methods for this problem (I did not ask for inference), but the methods I used was to get fitted values from the regression models and obtain descriptive statistics by treatment arm and sex.:*

Results: **Complete remissions were induced in 21 of 30 males (70.0%) on the IDR arm, and 17 of 35 males (48,6%) on the DNR arm. For females, complete remissions were induced in 30 of 35 females (85.7%) on the IDR arm, and 21 of 30 females (70.0%) on the DNR arm.**

We obtain similar results for all measures of treatment effect (RD, RR, and OR):

- **For the regressions used in problem 4 (a model including only an intercept and the main effects for treatment and sex), we fit a model with three parameters to a dataset that had four unique combinations of the covariate values included in the model. The fitted values for the proportions do not agree exactly with the sample proportions. (In this case of fitting an interaction model, the fitted values will tend to agree more closely with the sample proportions for those scales that seem to exhibit the least effect modification.**
- **On the other hand, in problem 7 (a model including the intercept, the main effects for treatment and sex, and their multiplicative interaction), we fit a model with four parameters to a dataset that had four unique combinations of the modeled covariate values (a "saturated" model: in this case, four equations in four unknowns). Thus the fitted values for each covariate combination are exactly equal to the corresponding sample proportions**

- **RD:The fitted values for the proportion of remission were**
  - **Males on IDR: Prob #4: 68.5% Prob #7: 70.0%**
  - **Males on DNR: Prob #4: 49.9% Prob #7: 48.6%**
  - **Females on IDR: Prob #4: 87.0% Prob #7: 85.7%**
  - **Females on DNR: Prob #4: 68.5% Prob #7: 70.0%**
- **OR:The fitted values for the proportion of remission were**
  - **Males on IDR: Prob #4: 70.2% Prob #7: 70.0%**
  - **Males on DNR: Prob #4: 48.4% Prob #7: 48.6%**
  - **Females on IDR: Prob #4: 85.5% Prob #7: 85.7%**
  - **Females on DNR: Prob #4: 70.2% Prob #7: 70.0%**
- **RD:The fitted values for the proportion of remission were**
  - **Males on IDR: Prob #4: 67.1% Prob #7: 70.0%**
  - **Males on DNR: Prob #4: 51.1% Prob #7: 48.6%**
  - **Females on IDR: Prob #4: 88.2% Prob #7: 85.7%**
  - **Females on DNR: Prob #4: 67.1% Prob #7: 70.0%**

10. Which of the above analyses should be used to decide whether idarubicin should be approved for the indication of AML? What problems exist with the use of the other analyses you performed?

<u>Instructions for grading</u>: This problem is worth 10 points. Any mention of choosing the analysis based on the p value should result in the student getting no credit for the problem. Otherwise, points should be assigned according to the student addressing at least two or more of the following points.

Ans: The major criterion for choosing the best analysis is what was pre-specified as the primary statistical analysis for the RCT.

Issues that might be onsidered when choosing the primary statistical analysis for a binary response variable include

- **For a relatively common outcome, the RD is the usual summarization of treatment effect for unadjusted analyses. This is in large part due to the greater interpretability of the RD with respect to clinical importance.**
- **For a rare outcome (such as in safety studies designed to investigate serious adverse events), the RR is the usual summarization of treatment effect. And an OR is approximately equal to the RR for rare events.**
- **When planning on adjustment for important prognostic variables (e.g., with stratified randomization), analyses based on OR might be selected in part due to the empirically observed tendency for their to be less effect modification.**
- **There was nothing in the description of this scientific problem to suggest that sex would be an important prognostic variable or an effect modifier, so choosing an unadjusted analysis might be the most straightforward.**
- **A model including an interaction should generally not be used unless 1) effect modification is a major part of the scientific question, or 2) it is anticipated that strong effect modification exists and modeling the interaction might provide greater precision. As a general rule this last condition will not be very much of a factor with logistic regression.**
- **As a general rule, including covariates that are not strongly prognostic will detract from the power of the study.**

Appendix: Stata Commands and Output

```
. * A Stata dataset was previously saved after formatting the leukemia dataset
. use z:documents/teach/datasets/leukemia

. describe

Contains data from z:documents/teach/datasets/leukemia.dta
  obs:           130
 vars:            30                          2 Mar 2010 16:12
 size:        17,550 (99.8% of memory free)
-------------------------------------------------------------------------------
              storage  display     value
variable name   type   format      label      variable label
-------------------------------------------------------------------------------
ptid            float  %9.0g
onstudy         str10  %10s
age             float  %9.0g
fab             float  %9.0g
karn            float  %9.0g
wbc             float  %9.0g
plt             float  %9.0g
hgb             float  %9.0g
crchemo         float  %9.0g
crdate          str10  %10s
fudate          str10  %10s
status          str2   %9s
bmtxdate        str10  %10s
donstudy        float  %9.0g
tx              float  %9.0g
male            float  %9.0g
eval            float  %9.0g
cr              float  %9.0g
dcrdate         float  %9.0g
dfudate         float  %9.0g
death           float  %9.0g
bmtx            float  %9.0g
dbmtxdate       float  %9.0g
incl            float  %9.0g
dthtime         float  %9.0g
cens            float  %9.0g
_st             byte   %8.0g
_d              byte   %8.0g
_t              int    %10.0g
```

```
_t0             byte   %10.0g
-------------------------------------------------------------------------------------------
Sorted by:  tx

.
. * Descriptive statistics used to answer Question #1
. tabstat age karn wbc plt hgb, col(stat) stat(n mean sd min q max) by(tx) long

tx           variable |         N      mean        sd       min       p25       p50       p75       max
----------------------+--------------------------------------------------------------------------------
0                 age |        65  39.78462  13.44839        19        27        40        54        60
                 karn |        65  79.53846  12.55373        40        80        80        90       100
                  wbc |        64  43.32344  55.01674        .7      2.85      16.7     69.85       215
                  plt |        64  93.57813  92.39285        11      35.5        62     128.5       457
                  hgb |        64    9.6375  1.489061       6.4      8.65      9.45     10.25      13.9
----------------------+--------------------------------------------------------------------------------
1                 age |        65        38  12.46746        17        27        36        49        61
                 karn |        65  79.53846  11.64994        30        70        80        90       100
                  wbc |        65  29.04308  36.31832        .4       2.9      11.8      41.9     154.1
                  plt |        65  66.58462  57.79433        11        32        50        78       370
                  hgb |        65  9.223077  1.821895       2.8         8       9.2      10.2      13.7
----------------------+--------------------------------------------------------------------------------
Total             age |       130  38.89231  12.94787        17        27      36.5        53        61
                 karn |       130  79.53846  12.06324        30        80        80        90       100
                  wbc |       129  36.12791   46.9112        .4       2.9      13.8      56.9       215
                  plt |       129  79.97674  77.81513        11        33        57        95       457
                  hgb |       129  9.428682  1.671601       2.8       8.3       9.3      10.2      13.9
-------------------------------------------------------------------------------------------

. tabulate male tx, col

+-------------------+
| Key               |
|-------------------|
|     frequency     |
| column percentage |
+-------------------+

           |          tx
      male |         0          1 |     Total
-----------+----------------------+----------
         0 |        30         35 |        65
           |     46.15      53.85 |     50.00
-----------+----------------------+----------
         1 |        35         30 |        65
           |     53.85      46.15 |     50.00
```

```
-----------+----------------------+----------
     Total |        65         65 |       130
           |    100.00     100.00 |    100.00


. tabulate fab tx, col

+-------------------+
| Key               |
|-------------------|
|     frequency     |
| column percentage |
+-------------------+

           |          tx
       fab |         0          1 |     Total
-----------+----------------------+----------
         0 |         1          0 |         1
           |      1.75       0.00 |      0.85
-----------+----------------------+----------
         1 |         6         13 |        19
           |     10.53      21.31 |     16.10
-----------+----------------------+----------
         2 |        15         15 |        30
           |     26.32      24.59 |     25.42
-----------+----------------------+----------
         3 |         9         11 |        20
           |     15.79      18.03 |     16.95
-----------+----------------------+----------
         4 |        12          8 |        20
           |     21.05      13.11 |     16.95
-----------+----------------------+----------
         5 |        13         12 |        25
           |     22.81      19.67 |     21.19
-----------+----------------------+----------
         6 |         1          2 |         3
           |      1.75       3.28 |      2.54
-----------+----------------------+----------
     Total |        57         61 |       118
           |    100.00     100.00 |    100.00


.
. * Analyses to answer Question #2
. * I use the Stata cs command, which provides inference for RD, RR, and OR
. * The chi square statistic is the default, I also execute it with the exact option
```

```
. * (I obtained the p values for unconditional exact tests using R)
. cs cr tx, or

                 | tx                      |
                 |   Exposed   Unexposed   |      Total
-----------------+-------------------------+------------
           Cases |        51          38   |         89
        Noncases |        14          27   |         41
-----------------+-------------------------+------------
           Total |        65          65   |        130
                 |                         |
            Risk |  .7846154    .5846154   |   .6846154
                 |                         |
                 |      Point estimate     |    [95% Conf. Interval]
                 |-------------------------+------------------------
 Risk difference |               .2        |    .0439899    .3560101
      Risk ratio |         1.342105        |    1.054389    1.708332
 Attr. frac. ex. |          .254902        |    .0515836    .4146336
 Attr. frac. pop |         .1460674        |
      Odds ratio |         2.588346        |     1.20652    5.545979 (Cornfield)
                 +-------------------------------------------------
                              chi2(1) =     6.02  Pr>chi2 = 0.0141

. cs cr tx, or exact

                 | tx                      |
                 |   Exposed   Unexposed   |      Total
-----------------+-------------------------+------------
           Cases |        51          38   |         89
        Noncases |        14          27   |         41
-----------------+-------------------------+------------
           Total |        65          65   |        130
                 |                         |
            Risk |  .7846154    .5846154   |   .6846154
                 |                         |
                 |      Point estimate     |    [95% Conf. Interval]
                 |-------------------------+------------------------
 Risk difference |               .2        |    .0439899    .3560101
      Risk ratio |         1.342105        |    1.054389    1.708332
 Attr. frac. ex. |          .254902        |    .0515836    .4146336
 Attr. frac. pop |         .1460674        |
      Odds ratio |         2.588346        |     1.20652    5.545979 (Cornfield)
                 +-------------------------------------------------
                                1-sided Fisher's exact P = 0.0115
                                2-sided Fisher's exact P = 0.0229
```

```
.
. * Analyses to answer Question #4
. * I consider both stratified and regression models for each of RD, RR, OR
. * First RD
. regress cr tx male, robust

Linear regression                                      Number of obs =     130
                                                       F(  2,   127) =    6.33
                                                       Prob > F      =  0.0024
                                                       R-squared     =  0.0860
                                                       Root MSE      =  .44945

------------------------------------------------------------------------------
             |               Robust
          cr |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          tx |   .1857143   .0794415     2.34   0.021     .0285138    .3429148
        male |  -.1857143   .0794415    -2.34   0.021    -.3429148   -.0285138
       _cons |   .6846154   .0716981     9.55   0.000     .5427378     .826493
------------------------------------------------------------------------------

. cs cr tx, by(male) istandard rd

            male |       RD       [95% Conf. Interval]       Weight
-----------------+------------------------------------------------
               0 |    .1571429    -.0436797   .3579655           35
               1 |    .2142857    -.0187525   .4473239           30
-----------------+------------------------------------------------
           Crude |          .2     .0439899   .3560101
 I. Standardized |    .1835165     .0309991   .3360338

.
. * Second OR
. logistic cr tx male

Logistic regression                               Number of obs   =        130
                                                  LR chi2(2)      =      11.56
                                                  Prob > chi2     =     0.0031
Log likelihood = -75.252997                       Pseudo R2       =     0.0713

------------------------------------------------------------------------------
          cr | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          tx |   2.511585   1.007479     2.30   0.022     1.144203    5.513062
        male |   .3981549    .159713    -2.30   0.022     .1813874    .8739708
------------------------------------------------------------------------------
```

```
. cs cr tx, by(male) or

            male |       OR       [95% Conf. Interval]   M-H Weight
-----------------+-------------------------------------------------
               0 |    2.571429     .7807859   8.404801     1.615385 (Cornfield)
               1 |    2.470588     .8981522   6.780976     2.353846 (Cornfield)
-----------------+-------------------------------------------------
           Crude |    2.588346      1.20652   5.545979
    M-H combined |    2.511628     1.144075   5.513866
-----------------------------------------------------------------
Test of homogeneity (M-H)      chi2(1) =    0.002  Pr>chi2 = 0.9609

                   Test that combined OR = 1:
                               Mantel-Haenszel chi2(1) =      5.33
                                                Pr>chi2 =    0.0210


.
. * Third RR (with an option to print the RR estimates instead of log RR)
. poisson cr tx male, ir

Iteration 0:   log likelihood = -120.95197
Iteration 1:   log likelihood = -120.95197

Poisson regression                                Number of obs   =        130
                                                  LR chi2(2)      =       3.54
                                                  Prob > chi2     =     0.1703
Log likelihood = -120.95197                       Pseudo R2       =     0.0144

------------------------------------------------------------------------------
          cr |        IRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          tx |    1.31435   .2824691     1.27   0.203     .8625375     2.00283
        male |   .7608323   .1635117    -1.27   0.203     .4992934     1.15937
------------------------------------------------------------------------------

. cs cr tx, by(male)

            male |       RR       [95% Conf. Interval]   M-H Weight
-----------------+-------------------------------------------------
               0 |     1.22449     .9342825   1.604841     11.30769
               1 |    1.441176     .9529711   2.179489     7.846154
-----------------+-------------------------------------------------
           Crude |    1.342105     1.054389   1.708332
    M-H combined |    1.313253     1.038088   1.661355
-----------------------------------------------------------------
```

```
Test of homogeneity (M-H)      chi2(1) =    0.451  Pr>chi2 = 0.5018

.
. * Analyses to answer Question #5
. * (Analyses similar to those in Question #2, but restricted to males)
. cs cr tx if male==1, or

                 | tx                     |
                 |   Exposed   Unexposed  |      Total
-----------------+------------------------+------------
           Cases |        21          17  |         38
        Noncases |         9          18  |         27
-----------------+------------------------+------------
           Total |        30          35  |         65
                 |                        |
            Risk |        .7    .4857143  |   .5846154
                 |                        |
                 |      Point estimate    |    [95% Conf. Interval]
                 |------------------------+------------------------
 Risk difference |         .2142857       |   -.0187525    .4473239
      Risk ratio |         1.441176       |    .9529711    2.179489
 Attr. frac. ex. |         .3061224       |   -.0493498    .5411768
 Attr. frac. pop |         .1691729       |
      Odds ratio |         2.470588       |    .8981522    6.780976 (Cornfield)
                 +-------------------------------------------------
                               chi2(1) =     3.05  Pr>chi2 = 0.0805

. cs cr tx if male==1, or exact

                 | tx                     |
                 |   Exposed   Unexposed  |      Total
-----------------+------------------------+------------
           Cases |        21          17  |         38
        Noncases |         9          18  |         27
-----------------+------------------------+------------
           Total |        30          35  |         65
                 |                        |
            Risk |        .7    .4857143  |   .5846154
                 |                        |
                 |      Point estimate    |    [95% Conf. Interval]
                 |------------------------+------------------------
 Risk difference |         .2142857       |   -.0187525    .4473239
      Risk ratio |         1.441176       |    .9529711    2.179489
 Attr. frac. ex. |         .3061224       |   -.0493498    .5411768
 Attr. frac. pop |         .1691729       |
      Odds ratio |         2.470588       |    .8981522    6.780976 (Cornfield)
```

```
                  +-------------------------------------------------
                                   1-sided Fisher's exact P = 0.0669
                                   2-sided Fisher's exact P = 0.1293

.
. * Analyses to answer Question #6
. * (Analyses similar to those in Question #2, but restricted to females)
. cs cr tx if male==0, or

                  | tx                      |
                  |   Exposed   Unexposed   |      Total
-----------------+------------------------+------------
           Cases |        30          21   |         51
        Noncases |         5           9   |         14
-----------------+------------------------+------------
           Total |        35          30   |         65
                 |                         |
            Risk |  .8571429          .7   |   .7846154
                 |                         |
                 |      Point estimate     |    [95% Conf. Interval]
                 |-------------------------+------------------------
 Risk difference |         .1571429        |   -.0436797    .3579655
      Risk ratio |          1.22449        |    .9342825    1.604841
 Attr. frac. ex. |         .1833333        |   -.0703401    .3768855
 Attr. frac. pop |         .1078431        |
      Odds ratio |         2.571429        |    .7807859    8.404801 (Cornfield)
                 +-------------------------------------------------
                                chi2(1) =     2.36  Pr>chi2 = 0.1244

. cs cr tx if male==0, or exact

                  | tx                      |
                  |   Exposed   Unexposed   |      Total
-----------------+------------------------+------------
           Cases |        30          21   |         51
        Noncases |         5           9   |         14
-----------------+------------------------+------------
           Total |        35          30   |         65
                 |                         |
            Risk |  .8571429          .7   |   .7846154
                 |                         |
                 |      Point estimate     |    [95% Conf. Interval]
                 |-------------------------+------------------------
 Risk difference |         .1571429        |   -.0436797    .3579655
      Risk ratio |          1.22449        |    .9342825    1.604841
 Attr. frac. ex. |         .1833333        |   -.0703401    .3768855
```

```
 Attr. frac. pop |         .1078431       |
      Odds ratio |         2.571429       |    .7807859    8.404801 (Cornfield)
                 +-------------------------------------------------
                                  1-sided Fisher's exact P = 0.1087
                                  2-sided Fisher's exact P = 0.1432

.
. * Analyses to answer Question #7
. * I consider both regression models, as well as gleaning the tests of homogeneity
. * from the stratified analyses used in Question #4
. regress cr tx male i.male##i.tx, robust
note: 1.male omitted because of collinearity
note: 1.tx omitted because of collinearity

Linear regression                                      Number of obs =     130
                                                       F(  3,   126) =    4.26
                                                       Prob > F      =  0.0067
                                                       R-squared     =  0.0870
                                                       Root MSE      =    .451

------------------------------------------------------------------------------
             |               Robust
          cr |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          tx |   .1571429   .1040761     1.51   0.134    -.0488206    .3631063
        male |  -.2142857   .1207718    -1.77   0.078    -.4532895    .0247181
      1.male |  (omitted)
        1.tx |  (omitted)
             |
     male#tx |
        1 1  |   .0571429   .1594291     0.36   0.721    -.2583627    .3726484
             |
       _cons |         .7   .0849837     8.24   0.000     .5318198    .8681802
------------------------------------------------------------------------------

. logistic cr tx male i.male##i.tx
note: 1.male omitted because of collinearity
note: 1.tx omitted because of collinearity

Logistic regression                               Number of obs   =        130
                                                  LR chi2(3)      =      11.57
                                                  Prob > chi2     =     0.0090
Log likelihood = -75.251793                       Pseudo R2       =     0.0714

------------------------------------------------------------------------------
          cr | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
```

```
-------------+----------------------------------------------------------------
          tx |   2.571429   1.610099     1.51   0.131     .7536936     8.77312
        male |   .4047619    .211528    -1.73   0.084     .1453326    1.127292
      1.male |  (omitted)
        1.tx |  (omitted)
             |
     male#tx |
        1 1  |   .9607843   .7835976    -0.05   0.961     .1942683     4.75171
------------------------------------------------------------------------------

. poisson cr tx male i.male##i.tx, robust ir
note: 1.male omitted because of collinearity
note: 1.tx omitted because of collinearity

Iteration 0:   log pseudolikelihood = -120.88116
Iteration 1:   log pseudolikelihood = -120.88116

Poisson regression                                      Number of obs   =        130
                                                        Wald chi2(3)    =      10.53
                                                        Prob > chi2     =     0.0146
Log pseudolikelihood = -120.88116                       Pseudo R2       =     0.0150

------------------------------------------------------------------------------
             |               Robust
          cr |        IRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          tx |    1.22449   .1696494     1.46   0.144     .9333053    1.606522
        male |   .6938776   .1470023    -1.73   0.085     .4580897     1.05103
      1.male |  (omitted)
        1.tx |  (omitted)
             |
     male#tx |
        1 1  |   1.176961   .2979323     0.64   0.520      .716624    1.933003
------------------------------------------------------------------------------

.
. * Analyses to answer Question #8
. * I re-execute the models in order to use the post-estimation testparm commands
. regress cr tx male i.male##i.tx, robust
note: 1.male omitted because of collinearity
note: 1.tx omitted because of collinearity

Linear regression                                      Number of obs =     130
                                                       F(  3,   126) =    4.26
                                                       Prob > F      =  0.0067
                                                       R-squared     =  0.0870
```

```
                                                      Root MSE      =    .451

------------------------------------------------------------------------------
             |               Robust
          cr |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          tx |   .1571429   .1040761     1.51   0.134    -.0488206    .3631063
        male |  -.2142857   .1207718    -1.77   0.078    -.4532895    .0247181
      1.male |  (omitted)
        1.tx |  (omitted)
             |
     male#tx |
        1 1  |   .0571429   .1594291     0.36   0.721    -.2583627    .3726484
             |
       _cons |         .7   .0849837     8.24   0.000     .5318198    .8681802
------------------------------------------------------------------------------

. testparm tx i.male#i.tx

 ( 1)  tx = 0
 ( 2)  1.male#1.tx = 0

       F(  2,   126) =    2.71
            Prob > F =    0.0701

.
. logistic cr tx male i.male##i.tx
note: 1.male omitted because of collinearity
note: 1.tx omitted because of collinearity

Logistic regression                               Number of obs   =        130
                                                  LR chi2(3)      =      11.57
                                                  Prob > chi2     =     0.0090
Log likelihood = -75.251793                       Pseudo R2       =     0.0714

------------------------------------------------------------------------------
          cr | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          tx |   2.571429   1.610099     1.51   0.131     .7536936     8.77312
        male |   .4047619    .211528    -1.73   0.084     .1453326    1.127292
      1.male |  (omitted)
        1.tx |  (omitted)
             |
     male#tx |
        1 1  |   .9607843   .7835976    -0.05   0.961     .1942683     4.75171
------------------------------------------------------------------------------
```

```
. testparm tx i.male#i.tx

 ( 1)  [cr]tx = 0
 ( 2)  [cr]1.male#1.tx = 0

           chi2(  2) =    5.27
         Prob > chi2 =    0.0717

.
. poisson cr tx male i.male##i.tx, robust ir
note: 1.male omitted because of collinearity
note: 1.tx omitted because of collinearity

Iteration 0:   log pseudolikelihood = -120.88116
Iteration 1:   log pseudolikelihood = -120.88116

Poisson regression                                      Number of obs   =        130
                                                        Wald chi2(3)    =      10.53
                                                        Prob > chi2     =     0.0146
Log pseudolikelihood = -120.88116                       Pseudo R2       =     0.0150

------------------------------------------------------------------------------
             |               Robust
          cr |        IRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          tx |    1.22449   .1696494     1.46   0.144     .9333053    1.606522
        male |   .6938776   .1470023    -1.73   0.085     .4580897     1.05103
      1.male |  (omitted)
        1.tx |  (omitted)
             |
     male#tx |
        1 1  |   1.176961   .2979323     0.64   0.520      .716624    1.933003
------------------------------------------------------------------------------

. testparm tx i.male#i.tx

 ( 1)  [cr]tx = 0
 ( 2)  [cr]1.male#1.tx = 0

           chi2(  2) =    5.11
         Prob > chi2 =    0.0776

.
. * Analyses to answer Question #9
. * I re-execute the models from Questions #4 and #7 in order to save the fitted values
```

```
. * I then compute summary statistics
. regress cr tx male

      Source |       SS       df       MS              Number of obs =     130
-------------+------------------------------           F(  2,   127) =    5.98
       Model |  2.41428571     2  1.20714286           Prob > F      =  0.0033
    Residual |  25.6549451   127  .202007441           R-squared     =  0.0860
-------------+------------------------------           Adj R-squared =  0.0716
       Total |  28.0692308   129  .217590936           Root MSE      =  .44945

------------------------------------------------------------------------------
          cr |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          tx |   .1857143   .0790735     2.35   0.020     .0292422    .3421864
        male |  -.1857143   .0790735    -2.35   0.020    -.3421864   -.0292422
       _cons |   .6846154   .0701477     9.76   0.000     .5458058    .8234249
------------------------------------------------------------------------------

. predict fit
(option xb assumed; fitted values)

. bysort tx male: summ fit

-------------------------------------------------------------------------------------------------------
-> tx = 0, male = 0

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
         fit |        30    .6846154           0   .6846154   .6846154

-------------------------------------------------------------------------------------------------------
-> tx = 0, male = 1

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
         fit |        35    .4989011           0   .4989011   .4989011

-------------------------------------------------------------------------------------------------------
-> tx = 1, male = 0

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
         fit |        35    .8703297           0   .8703297   .8703297

-------------------------------------------------------------------------------------------------------
-> tx = 1, male = 1
```

```
    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
         fit |        30    .6846154           0   .6846154   .6846154


. drop fit

.
. regress cr tx male i.male##i.tx
note: 1.male omitted because of collinearity
note: 1.tx omitted because of collinearity

      Source |       SS       df       MS              Number of obs =     130
-------------+------------------------------           F(  3,   126) =    4.00
       Model |  2.44065934     3  .813553114           Prob > F      =  0.0093
    Residual |  25.6285714   126  .203401361           R-squared     =  0.0870
-------------+------------------------------           Adj R-squared =  0.0652
       Total |  28.0692308   129  .217590936           Root MSE      =    .451

------------------------------------------------------------------------------
          cr |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          tx |   .1571429   .1122119     1.40   0.164    -.0649212    .3792069
        male |  -.2142857   .1122119    -1.91   0.058    -.4363498    .0077784
      1.male |  (omitted)
        1.tx |  (omitted)
             |
     male#tx |
        1 1  |   .0571429   .1586916     0.36   0.719    -.2569032    .3711889
             |
       _cons |         .7    .082341     8.50   0.000     .5370495    .8629505
------------------------------------------------------------------------------

. predict fit
(option xb assumed; fitted values)

. bysort tx male: summ fit

-------------------------------------------------------------------------------------------------------------
-> tx = 0, male = 0

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
         fit |        30          .7           0         .7         .7
```

```
-------------------------------------------------------------------------------------------------------
-> tx = 0, male = 1

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
         fit |        35    .4857143           0   .4857143   .4857143

-------------------------------------------------------------------------------------------------------
-> tx = 1, male = 0

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
         fit |        35    .8571429           0   .8571429   .8571429

-------------------------------------------------------------------------------------------------------
-> tx = 1, male = 1

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
         fit |        30          .7           0         .7         .7


. drop fit

.
. logistic cr tx male

Logistic regression                               Number of obs   =        130
                                                  LR chi2(2)      =      11.56
                                                  Prob > chi2     =     0.0031
Log likelihood = -75.252997                       Pseudo R2       =     0.0713

------------------------------------------------------------------------------
          cr | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          tx |   2.511585   1.007479     2.30   0.022     1.144203    5.513062
        male |   .3981549    .159713    -2.30   0.022     .1813874    .8739708
------------------------------------------------------------------------------

. predict fit
(option pr assumed; Pr(cr))

. bysort tx male: summ fit

-------------------------------------------------------------------------------------------------------
-> tx = 0, male = 0
```

```
    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
         fit |        30    .7020064           0   .7020064   .7020064

-------------------------------------------------------------------------------------------------------------
-> tx = 0, male = 1

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
         fit |        35    .4839945           0   .4839945   .4839945

-------------------------------------------------------------------------------------------------------------
-> tx = 1, male = 0

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
         fit |        35    .8554231           0   .8554231   .8554231

-------------------------------------------------------------------------------------------------------------
-> tx = 1, male = 1

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
         fit |        30    .7020064           0   .7020064   .7020064


. drop fit

.
. logistic cr tx male i.male##i.tx
note: 1.male omitted because of collinearity
note: 1.tx omitted because of collinearity

Logistic regression                               Number of obs   =        130
                                                  LR chi2(3)      =      11.57
                                                  Prob > chi2     =     0.0090
Log likelihood = -75.251793                       Pseudo R2       =     0.0714

------------------------------------------------------------------------------
          cr | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          tx |   2.571429   1.610099     1.51   0.131     .7536936     8.77312
        male |   .4047619    .211528    -1.73   0.084     .1453326    1.127292
      1.male |  (omitted)
        1.tx |  (omitted)
```

```
             |
     male#tx |
        1 1  |   .9607843   .7835976    -0.05   0.961     .1942683     4.75171
------------------------------------------------------------------------------

. predict fit
(option pr assumed; Pr(cr))

. bysort tx male: summ fit

-------------------------------------------------------------------------------------------------------
-> tx = 0, male = 0

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
         fit |        30          .7           0         .7         .7

-------------------------------------------------------------------------------------------------------
-> tx = 0, male = 1

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
         fit |        35    .4857143           0   .4857143   .4857143

-------------------------------------------------------------------------------------------------------
-> tx = 1, male = 0

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
         fit |        35    .8571429           0   .8571429   .8571429

-------------------------------------------------------------------------------------------------------
-> tx = 1, male = 1

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
         fit |        30          .7           0         .7         .7


. drop fit

.
. poisson cr tx male

Iteration 0:   log likelihood = -120.95197
Iteration 1:   log likelihood = -120.95197
```

```
Poisson regression                                Number of obs   =        130
                                                  LR chi2(2)      =       3.54
                                                  Prob > chi2     =     0.1703
Log likelihood = -120.95197                       Pseudo R2       =     0.0144

------------------------------------------------------------------------------
          cr |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          tx |   .2733423   .2149116     1.27   0.203    -.1478767    .6945613
        male |  -.2733423   .2149116    -1.27   0.203    -.6945613    .1478767
       _cons |  -.3989374   .1911251    -2.09   0.037    -.7735357   -.0243391
------------------------------------------------------------------------------

. predict fit
(option n assumed; predicted number of events)

. bysort tx male: summ fit

-------------------------------------------------------------------------------------------------------
-> tx = 0, male = 0

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
         fit |        30    .6710327           0   .6710327   .6710327

-------------------------------------------------------------------------------------------------------
-> tx = 0, male = 1

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
         fit |        35    .5105433           0   .5105433   .5105433

-------------------------------------------------------------------------------------------------------
-> tx = 1, male = 0

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
         fit |        35     .881972           0    .881972    .881972

-------------------------------------------------------------------------------------------------------
-> tx = 1, male = 1

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
         fit |        30    .6710327           0   .6710327   .6710327
```

```
. drop fit

.
. poisson cr tx male i.male##i.tx
note: 1.male omitted because of collinearity
note: 1.tx omitted because of collinearity

Iteration 0:   log likelihood = -120.88116
Iteration 1:   log likelihood = -120.88116

Poisson regression                                Number of obs   =        130
                                                  LR chi2(3)      =       3.68
                                                  Prob > chi2     =     0.2980
Log likelihood = -120.88116                       Pseudo R2       =     0.0150

------------------------------------------------------------------------------
          cr |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          tx |   .2025242   .2845213     0.71   0.477    -.3551273    .7601758
        male |  -.3654596   .3262554    -1.12   0.263    -1.004908    .2739891
      1.male |  (omitted)
        1.tx |  (omitted)
             |
     male#tx |
        1 1  |   .1629354   .4328914     0.38   0.707    -.6855161    1.011387
             |
       _cons |  -.3566749   .2182179    -1.63   0.102    -.7843741    .0710243
------------------------------------------------------------------------------

. predict fit
(option n assumed; predicted number of events)

. bysort tx male: summ fit

-----------------------------------------------------------------------------------------------------------
-> tx = 0, male = 0

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
         fit |        30          .7           0         .7         .7

-----------------------------------------------------------------------------------------------------------
-> tx = 0, male = 1
```

```
    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
         fit |        35    .4857143           0   .4857143   .4857143

-----------------------------------------------------------------------------------------------------------------------
-> tx = 1, male = 0

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
         fit |        35    .8571429           0   .8571429   .8571429

-----------------------------------------------------------------------------------------------------------------------
-> tx = 1, male = 1

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
         fit |        30          .7           0         .7         .7


. drop fit
```