

Short solutions to the exam in HMET 5140

Exercise 1.

```
. twoway (scatter residual predicted)
```

```
. ttest costs, by(ward) unequal
```

Two-sample t test with unequal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
Urban	70	21456.97	824.718	6900.086	19811.7	23102.24
Rural	50	19258.98	455.9348	3223.946	18342.75	20175.22
combined	120	20541.14	524.9635	5750.687	19501.66	21580.62
diff		2197.987	942.3568		329.2472	4066.727

diff = mean(Urban) - mean(Rural) t = 2.3324
Ho: diff = 0 Satterthwaite's degrees of freedom = 103.949

Ha: diff < 0 Ha: diff != 0 Ha: diff > 0
Pr(T < t) = 0.9892 Pr(|T| > |t|) = 0.0216 Pr(T > t) = 0.0108

a) (8p) Null hypothesis: Mean costs in urban and rural ward are equal ($\mu_{\text{urban}} = \mu_{\text{rural}}$).

Alternative: They are different ($\mu_{\text{urban}} \neq \mu_{\text{rural}}$) (2p). Mean difference is 2197 kroner, with urban ward having the highest costs (2p). P-value is 2.2%, hence there is a difference in mean costs (2p). T-test with unequal variances are used is because the standard error of the urban group is close to twice as large as for the rural group (2p).

```
. ranksum costs, by(ward)
```

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

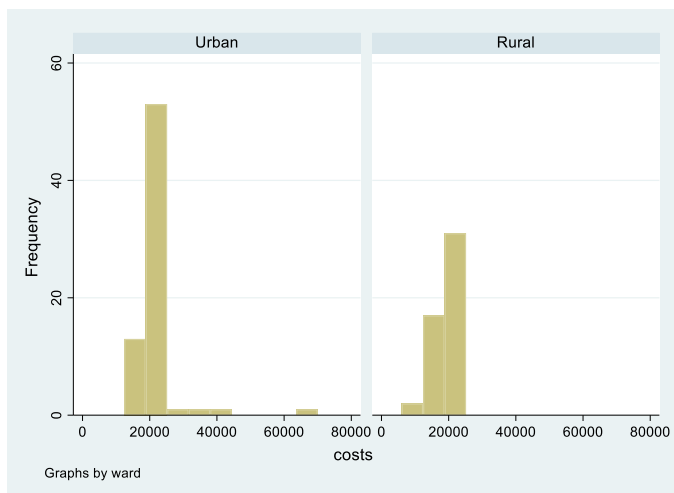
ward	obs	rank sum	expected
Urban	70	4590	4235
Rural	50	2670	3025
combined	120	7260	7260

unadjusted variance 35291.67
adjustment for ties 0.00

adjusted variance 35291.67

Ho: costs(ward==Urban) = costs(ward==Rural)

z = 1.890
Prob > |z| = 0.0588
Exact Prob = 0.0590



b) (8p) Null hypothesis: Median costs are equal in rural and urban wards (or more precisely, the distributions are equal). Alternative: The medians/distributions are different (2p). P-value for this test is 5.9%, hence there is no difference (2p).

We see from the histograms that the urban ward includes a few patients with very high costs, thus giving a skewed distribution. Apart from these observations, the distributions are similar (2p). Hence, when mean costs is considered, the mean at the urban ward is influenced by these few extreme observations to such an extent that the t-test gives the result of significant difference. When the influence of these observations is made smaller, as it is in the Wilcoxon test using ranks of observations instead of actual values, there is no significant difference anymore (2p).

```
. estat bootstrap, p
```

```
Bootstrap results
Number of strata   =           2           Number of obs   =          120
Replications      =          10000
```

```
command: ttest costs, by(hospital)
       _bs_1: r(mu_1)-r(mu_2)
```

	Observed Coef.	Bias	Bootstrap Std. Err.	[95% Conf. Interval]	
_bs_1	2197.987	2.604626	931.56507	602.7515	4200.881 (P)

(P) percentile confidence interval

c) (7p) Stratified means that data are bootstrapped within each treatment group separately, hence every bootstrap sample will always include 70 urban and 50 rural patients, same as the original sample (probably these sample sizes were arrived at by design, and not by coincidence, hence stratified sampling makes sense). If sampling was done unstratified, the number of patients in each group would vary from bootstrap sample to bootstrap sample (2p). The program will sample 50 rural and 70 urban patients, at random with replacement, to create each new bootstrap sample (1p). The output shows that 10000 samples are created (1p). The mean difference in costs is calculated and stored for every sample, and ordered from the smallest to the largest difference (1p). Observations no. 250 and 9750 are selected as the lower and upper boundary for the percentile method interval (1p). The 95% confidence interval means that the difference in costs between the urban and rural wards is from 602 kroner to 4200 kroner, giving the same conclusion as the t-test in a): There seems to be a significant difference in the costs (2p).

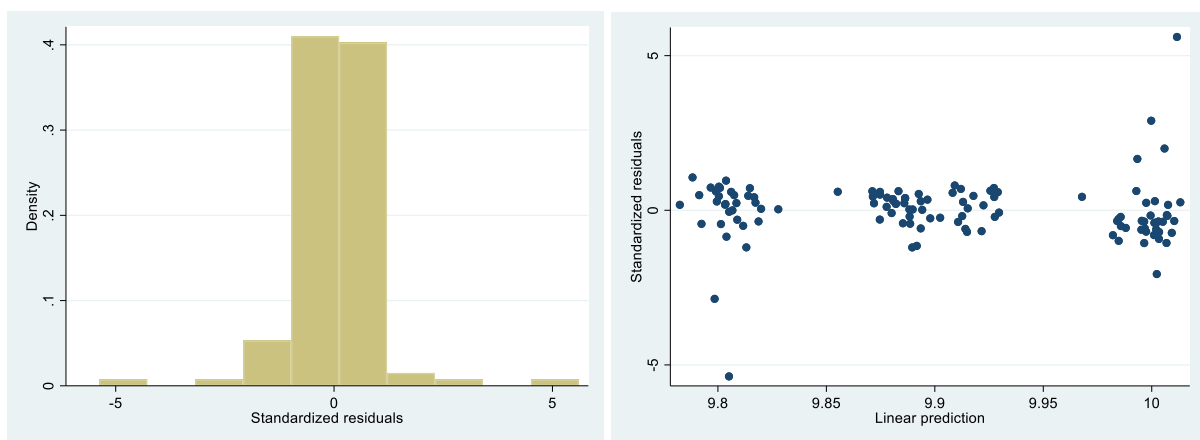
```
. regress logcosts ward age severity
```

Source	SS	df	MS	Number of obs	=	120
Model	.675576993	3	.225192331	F(3, 116)	=	5.19
Residual	5.03803306	116	.043431319	Prob > F	=	0.0021
				R-squared	=	0.1182
				Adj R-squared	=	0.0954
Total	5.71361005	119	.04801353	Root MSE	=	.2084

logcosts	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
ward	-.0615626	.0531598	-1.16	0.249	-.1668524	.0437271
age	.0009026	.0018538	0.49	0.627	-.002769	.0045742
severity	.1121738	.038925	2.88	0.005	.0350779	.1892698
_cons	9.812389	.1509746	64.99	0.000	9.513365	10.11141

d) (12p) The effect of high vs. low severity is $\exp(0.112)=1.12=12\%$ higher costs (2p). The confidence interval is $\exp(0.035)=1.04$, $\exp(0.189)=1.21$. As the interval does not contain 1, the effect is significant at 5% level (2p). The effect of reducing age by 20 years is $\exp(-20*0.0009)=0.982=1.8\%$ lower costs (non-significant). (2p)

The effect of rural vs. urban ward is $\exp(-0.062)=0.940=6\%$ lower costs (non-significant) (2p). We know from the point estimates that higher severity (significant) and age (non-significant) increase costs, and that the point estimate for the rural ward still indicates a lower costs vs. the urban ward, but with $p\text{-value} \gg 0.05$ (hence, less of a difference between urban and rural than previously in the exercise). This could then happen because the urban ward treats more patients with higher severity and/or age than the rural ward (4p). Any unspecific answer of effect of ward being confounded by age and/or severity will give 1p.



e) (4p) The residuals are supposed to be approximately normal if the log-linear model fit is ok. This looks not too bad in the left plot (2p). The right plot, however, describes variation around the model, which should be free of heteroscedasticity (i.e. be fairly constant for small and large predicted values). The variation seems to be constant except for just four observations that deviate from this trend (2p).

```
. tab2 severity ward
```

-> tabulation of severity by ward

severity	ward		Total
	Urban	Rural	
Low severity	32	32	64
High severity	38	18	56
Total	70	50	120

f) (8p) Hypotheses e.g.: $H_0: p_{\text{high severity, urban}} = p_{\text{high severity, rural}}$ $H_1: p_{\text{high severity, urban}} \neq p_{\text{high severity, rural}}$ (2p)

Expected frequencies:

$70*64/120=37.3$ $70*56/120=32.7$ $50*64/120=26.7$ $50*56/120=23.3$ (2p)

Test statistic: $\frac{(32-37.3)^2}{37.3} + \frac{(38-32.7)^2}{32.7} + \frac{(32-26.7)^2}{26.7} + \frac{(18-23.3)^2}{23.3} = 3.91$ (2p)

Should be compared to 95%-percentile of a chi-square distribution with $(2-1)*(2-1)=1$ d.f., i.e. 3.84 (1p). Reject H_0 since 3.91 is greater than 3.84. (1p)

```
Negative binomial regression      Number of obs   =      120
                                LR chi2(3)             =      11.26
Dispersion      = mean           Prob > chi2       =      0.0104
Log likelihood = -280.40061       Pseudo R2        =      0.0197
```

	los	IRR	Std. Err.	z	P> z	[95% Conf. Interval]
ward						
Rural		.7785349	.0978333	-1.99	0.046	.6085738 .9959623
severity						
High severity		1.255408	.1151896	2.48	0.013	1.048777 1.50275
age		.9961146	.0043061	-0.90	0.368	.9877104 1.00459
_cons		6.807601	2.380596	5.48	0.000	3.430296 13.51004
/lnalpha		-3.004312	.6807978			-4.338651 -1.669972
alpha		.0495729	.0337491			.0130541 .1882523

Note: Estimates are transformed only in the first equation.

Note: _cons estimates baseline incidence rate.

LR test of alpha=0: chibar2(01) = 2.94

Prob >= chibar2 = 0.043

g) (8p) The negative binomial model is suitable for LOS data because: LOS are usually strictly positive and skewed with a long right tail (1p), discrete outcome, so suitable for count type models (1p) and having greater variance than the mean, so a usually a better alternative than Poisson models (1p). The latter is precisely the meaning of the “alpha”-parameter in the table: This denotes the extra parameter in a neg bin model over a Poisson model, and as it is significantly greater than zero, one expects the neg bin model to provide better fit than the Poisson model for these data (2p). The effects are reported on the anti-log scale, and can be interpreted directly as incidence rate ratios. Severity is still significant, and having an effect in the same direction for LOS as for costs. Age is not significant, hence the oppsite sign of the coefficient does not matter that much. However, in this analysis with LOS as the outcome, ward is still borderline significant, with 22% lower LOS than the urban ward (3p).

```
. margins ward
```

```
Predictive margins      Number of obs   =      120
Model VCE      : OIM

Expression      : Predicted number of events, predict()
```

		Delta-method				
	Margin	Std. Err.	z	P> z	[95% Conf. Interval]	
ward						
Urban	5.751271	.3883992	14.81	0.000	4.990022 6.512519	
Rural	4.477565	.3932975	11.38	0.000	3.706716 5.248414	

```
. margins, dydx(ward)
```

```
Average marginal effects      Number of obs   =      120
Model VCE      : OIM

Expression      : Predicted number of events, predict()
dy/dx w.r.t. : 1.ward
```

		Delta-method				
	dy/dx	Std. Err.	z	P> z	[95% Conf. Interval]	
ward						
Rural	-1.273706	.6293324	-2.02	0.043	-2.507175 -.040237	

Note: dy/dx for factor levels is the discrete change from the base level.

h) (5p) A marginal effect means that the software uses the model in g) to calculate (predict from the model) what the average LOS would be if all patients in the data had stayed in the urban ward and values of variables age and severity as observed in the data, and the average LOS if all patients had stayed in the rural ward and values of age and severity as observed. Looking at the difference between the two, would be the MARGINAL effect of ward on LOS, that is, adjusted for other variables in model (3p). The output means that, when comparing patients of the same age and severity, the independent effect of urban vs rural ward is around 1.27 days shorter LOS in the rural ward (borderline significant from the neg bin model in g and also the confidence interval in second table above) (2p).

Grading:

0-24p=F

25-29p=E

29-35p=D

36-44p=C

45-54p=B

55-60p=A