******

**Answers Key Exercises - blavaan**

**Exercise 1 - Simple regression analysis**

**1-a1.** Both X1 and X2 are significant predictors of Y1. They are both positively related to Y1. X1 explains more of the variance in Y1, than does X2.

| **Coefficientsa** |
| --- |
| Model | Unstandardized Coefficients | Standardized Coefficients | t | Sig. | 95% Confidence Interval for B |
|  | B | Std. Error | Beta |  |  | Lower Bound | Upper Bound |
| 1 | (Constant) | .511 | .044 |  | 11.730 | .000 | .425 | .597 |
|  | X1 | .969 | .042 | .653 | 23.287 | .000 | .888 | 1.051 |
|  | X2 | .649 | .045 | .409 | 14.583 | .000 | .562 | .736 |
| a. Dependent Variable: Y1 |  |  |  |  |  |  |

**1-a2.** Both age and anxiety can be used to predict depression. Their relation to depression is significantly different from 0. The results in the ANOVA table are: *F*(2,497) = 387.13, p-value <0.001.

**1-b.** The results are almost the same as in SPSS:

Regressions:

 Estimate Std.Err z-value P(>|z|) ci.lower ci.upper

 y ~

 x1 0.969 0.042 23.357 0.000 0.888 1.051

 x2 0.649 0.044 14.627 0.000 0.562 0.736

Intercepts:

 Estimate Std.Err z-value P(>|z|) ci.lower ci.upper

 .y 0.511 0.043 11.766 0.000 0.426 0.596

Variances:

 Estimate Std.Err z-value P(>|z|) ci.lower ci.upper

 .y 0.941 0.060 15.811 0.000 0.825 1.058

R-Square:

 Estimate

 y 0.609

**1-c1.** The results of the Bayesian regression analysis with blavaan:

Note: Your results should be approximately the same. Results with Bayesian analysis can vary somewhat, therefore it don’t have to be exactly the same.

Regressions:

 Estimate Post.SD HPD.025 HPD.975 PSRF Prior

 y ~

 x1 0.969 0.042 0.886 1.048 1.000 dnorm(0,1e-2)

 x2 0.649 0.045 0.563 0.737 1.000 dnorm(0,1e-2)

Intercepts:

 Estimate Post.SD HPD.025 HPD.975 PSRF Prior

 .y 0.511 0.043 0.423 0.592 1.000 dnorm(0,1e-3)

Variances:

 Estimate Post.SD HPD.025 HPD.975 PSRF Prior

 .y 0.949 0.061 0.833 1.069 1.000 dgamma(1,.5)

R-Square:

 Estimate

 y 0.607

**1-c2**. The numeric results do not differ much from the frequentist analysis results. But the interpretation is a little different. We conclude that X1 and X2, age and anxiety, are both predictive of Y1, depression. The B estimates are the median of the posterior distribution of these parameters, so they represent the most likely value of this parameter given our prior and our observed data. The 95% *Credibility* Interval represents the range of values that we are 95% certain contains the true parameter value, given our prior and our observed data.

**1-c3.** Plots in blavaan don’t have exactly the same name as the parameter estimates, therefore it is useful to find out which plot belongs to which parameter estimate. The plot with beta[1,2,1] on the Y-axis, is the one for the regression coefficient x1. The plot with beta[1,3,1] on the Y-axis, is the one for regression coefficient x2. The plot with psi[1,1,1] on the Y-axis is the one for the variance, and the alpha[1,1,1] belongs to the intercept.

**Exercise 2- Sensitivity analysis**

**The results in the table below are an indication of how your results approximately should look like, the exact values can differ.**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Prior mean and prior variance used: |  | Mean IQ score: |  | 95% C.I./C.C.I.: |  | Bias with regard to Default prior |
| ML | - |  | 102.000 |  | 95.426,108.574 |  | - |
| Default prior | Mean= 0, var= 1000 |  | 100.807 |  | 93.696,107.490 |  | - |
| Subjective Prior 1 | Mean=100, var=100 |  | 101.802 |  | 95.368,108.118 |  | 0.99% |
| Subjective Prior 2 | Mean=100, var=1 |  | 100.167 |  | 98.249,101.967 |  | -0.63% |
| Subjective Prior 3 | Mean=0, var=1 |  | 0.212 |  | -1.791, 2.160 |  | -99.79% |
| Subjective Prior 4 | Mean=0, var=100 |  | 30.312 |  | 4.258, 56.559 |  | -69.93% |
| Subjective Prior 5 | Mean=0, var=1000 |  | 100.852 |  | 94.046,107.684 |  | 0.04% |

* The default prior for the mean (Intercepts) is dnorm(0, 1000). It’s a normally distributed prior with a mean of 0 and a variance of 1000. However, because blavaan uses the precision, the prior is specified like this: dnorm(0,1e-3). This result is shown in the parameter estimates table in the last column. The variance of the mean prior is very large.
* Using a large variance of 100, with a prior mean of 100, we can see that the credibility interval of the posterior mean is slightly smaller than when using the default prior.
* Mis-specifying the mean, and using a large variance (100), is not enough to get a good result, a larger variance of 1000 is better.
* The bias percentages show that Subjective prior 1 and 2 do not bias the mean too much with regard to the default prior. With a mis-specified mean and a variance of 1 and 100, the bias is far above acceptable (generally speaking, a bias lower than +/- 10% is acceptable). Subjective prior 5 actually has lower bias with a mis-specified mean, but a very wide variance, than subjective prior 1 (perfect mean, but very small variance).
* The R code for steps 6-9 should look like this:

#STEP6: Specifying a new alternative prior for the mean IQ score.

fit.IQ4 <- blavaan(model.IQ1, data=data.IQ, dp = dpriors(nu= "dnorm(100, 1)"))

summary(fit.IQ4, ci=TRUE)

#STEP7: Specifying a new alternative prior for the mean IQ score.

fit.IQ5 <- blavaan(model.IQ1, data=data.IQ, dp = dpriors(nu= "dnorm(0, 1)"))

summary(fit.IQ5, ci=TRUE)

#STEP8: Specifying a new alternative prior for the mean IQ score.

fit.IQ6 <- blavaan(model.IQ1, data=data.IQ, dp = dpriors(nu= "dnorm(0, 0.01)"))

summary(fit.IQ6, ci=TRUE)

#STEP9: Specifying a new alternative prior for the mean IQ score.

fit.IQ7 <- blavaan(model.IQ1, data=data.IQ, dp = dpriors(nu= "dnorm(0, 0.001)"))

summary(fit.IQ7, ci=TRUE)

**Exercise 3 – Violin plot**

