Exercise 1
Estimate a 1-parametric logistic model in Stata and plot the ICC curves.
Do any items appear redundant?

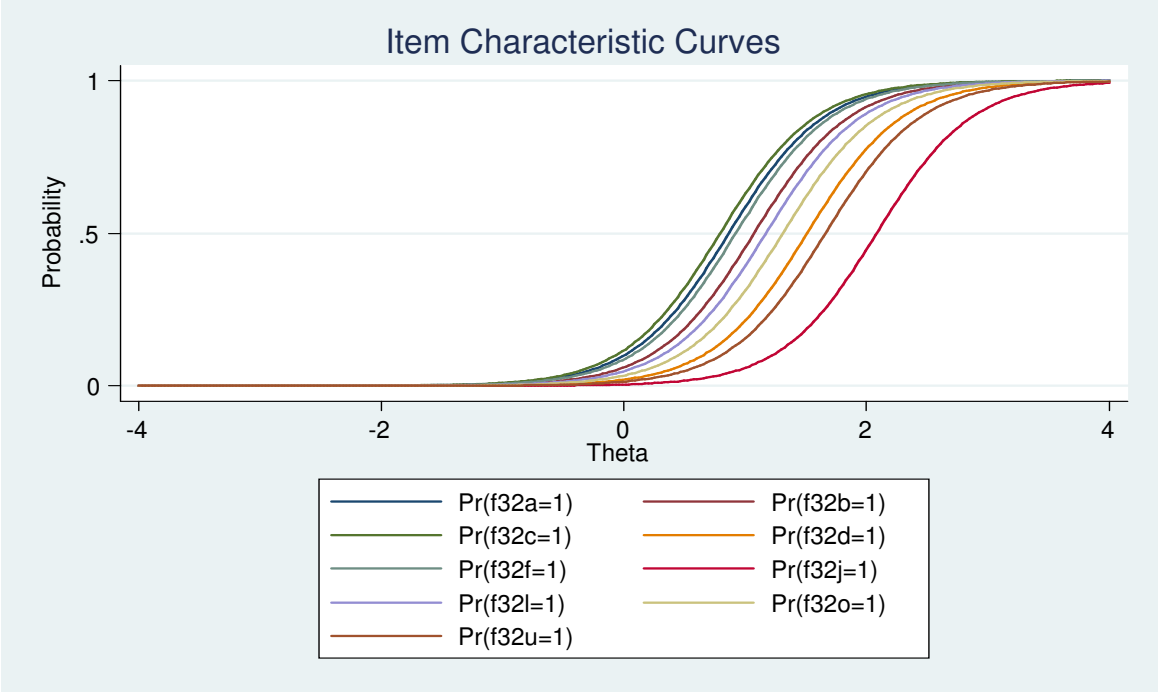
```
* 1pl - model
irt 1pl f32a f32b f32c f32d f32f f32j f32l f32o f32u

* the following commando list the parameters in a more compact form
estat report, byparm sort(b)
```

One-parameter logistic model — Number of obs = 482
Log likelihood = -1398.841

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Discrim	2.547997	.1691563	15.06	0.000	2.216457	2.879538
Diff						
f32c	.7928747	.073697	10.76	0.000	.6484313	.9373181
f32a	.8641884	.0758167	11.40	0.000	.7155904	1.012786
f32f	.9219419	.0776955	11.87	0.000	.7696616	1.074222
f32b	1.072519	.08325	12.88	0.000	.9093525	1.235686
f32l	1.169751	.0873257	13.40	0.000	.9985954	1.340906
f32o	1.309212	.0938285	13.95	0.000	1.125311	1.493112
f32d	1.510093	.1045656	14.44	0.000	1.305148	1.715038
f32u	1.660917	.1137402	14.60	0.000	1.43799	1.883843
f32j	2.083024	.1454028	14.33	0.000	1.79804	2.368008

```
* the ICC curves
irtgraph icc
```



The three items f32a, f32c, f32f have very similar difficulty parameters and they appear closest together in the graph. One of them might be redundant. Which one to drop depends on many other things: the actual wording of the items, the amount of missing, problems with different slope, DIF, and item information.

Exercise 2

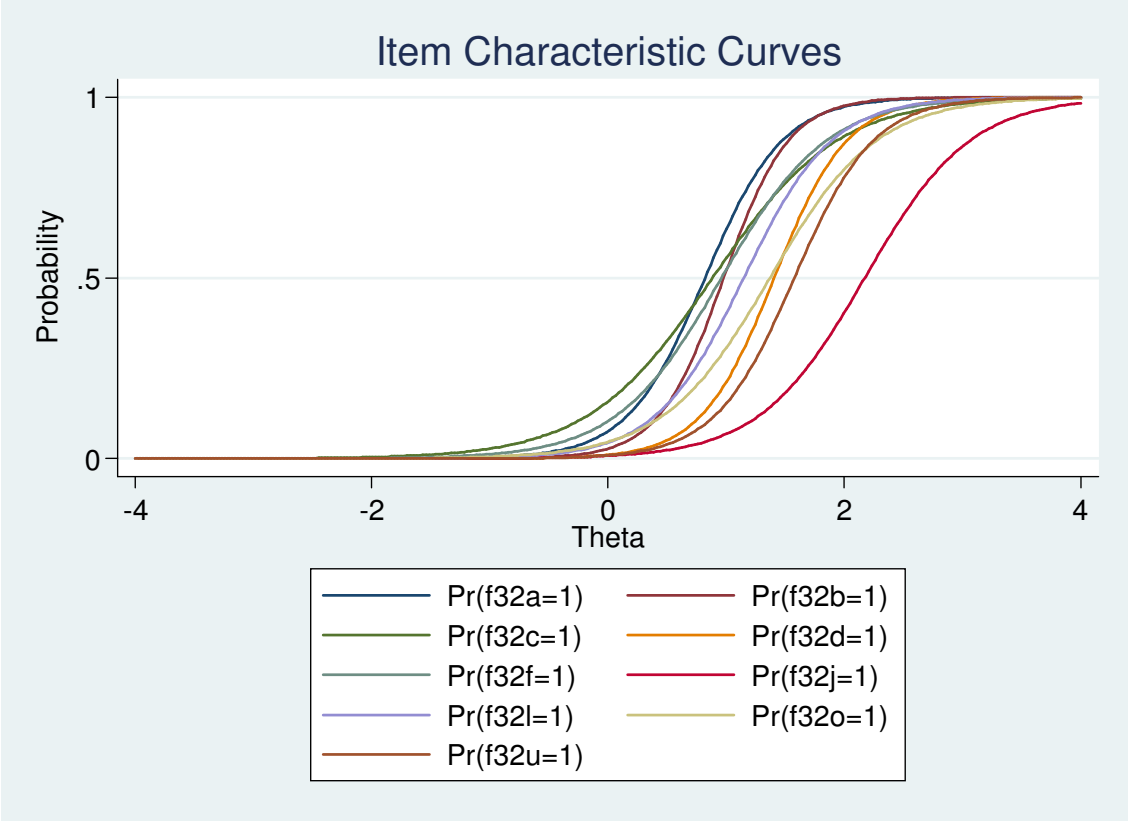
Estimate a 2-parameter logistic model in Stata and again plot the ICC curves.
Do the slopes differ?

```
* 2pl
irt 2pl f32a f32b f32c f32d f32f f32j f32l f32o f32u
estat report, byparm sort(b)
```

```
Two-parameter logistic model                    Number of obs     =        482
Log likelihood = -1389.0573
```

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Discrim						
f32a	3.050196	.4505825	6.77	0.000	2.16707	3.933321
f32c	1.891396	.2601937	7.27	0.000	1.381426	2.401367
f32f	2.245163	.3145108	7.14	0.000	1.628734	2.861593
f32b	3.643796	.5798689	6.28	0.000	2.507274	4.780318
f32l	2.676439	.3939481	6.79	0.000	1.904315	3.448563
f32o	2.213462	.332899	6.65	0.000	1.560992	2.865932
f32d	3.234467	.5538698	5.84	0.000	2.148902	4.320032
f32u	2.974973	.5254615	5.66	0.000	1.945087	4.004858
f32j	2.23679	.4524947	4.94	0.000	1.349917	3.123663
Diff						
f32a	.8218071	.074731	11.00	0.000	.675337	.9682772
f32c	.8852759	.0966496	9.16	0.000	.6958463	1.074706
f32f	.9614889	.092383	10.41	0.000	.7804215	1.142556
f32b	.9800225	.0770721	12.72	0.000	.8289639	1.131081
f32l	1.149778	.095882	11.99	0.000	.9618531	1.337704
f32o	1.3709	.122628	11.18	0.000	1.130554	1.611247
f32d	1.406876	.1068295	13.17	0.000	1.197494	1.616258
f32u	1.577335	.1259679	12.52	0.000	1.330442	1.824228
f32j	2.173269	.2315461	9.39	0.000	1.719447	2.627091

```
* the ICC curves
irtgraph icc
```



The smallest slope f32c=1.89 and the largest f32b=3.64. The graph does not reveal any items being obviously different but there are small variations.

Exercise 3
Test if the slopes can be equal for all items. Is a 1pl model adequate?

```
* first estimate each model and save the results via estimates store
* 1pl - model
irt 1pl f32a f32b f32c f32d f32f f32j f32l f32o f32u
estimates store model_1pl

* 2pl
irt 2pl f32a f32b f32c f32d f32f f32j f32l f32o f32u
estimates store model_2pl

* LR-test: does the data allow using the simpler 1pl model instead of the more
* complicated 2pl model?
lrtest model_1pl model_2pl
```

```
Likelihood-ratio test                                 LR chi2(8)  =     19.57
(Assumption: model_1pl nested in model_2pl)           Prob > chi2 =    0.0121
```

No - data does not allow that: p-value=0.0121. The best model so far is the 2pl model! The test tests whether it is ok to assume one common slope for all items or not.

Exercise 4
Explore whether there is DIF for age above or below 55 years on any for the items using the Mantel-Haenszel method.

```
* create variable indicating under or above 55 years of age
egen age55=cut(age), at(18,55,100) label

* DIF investigated via MH test
difmh f32a f32b f32c f32d f32f f32j f32l f32o f32u ///
    , group(age55)
```

```
Mantel-Haenszel DIF Analysis
```

Item	Chi2	Prob.	Odds Ratio	[95% Conf.	Interval]
f32a	1.56	0.2109	0.6175	0.3190	1.1956
f32b	3.95	0.0468	0.4283	0.1988	0.9225
f32c	0.33	0.5648	1.2336	0.6954	2.1882
f32d	1.89	0.1693	2.0262	0.8439	4.8649
f32f	0.37	0.5453	0.7945	0.4360	1.4479
f32j	4.02	0.0449	4.1044	1.1434	14.7325
f32l	0.68	0.4103	0.7026	0.3521	1.4019
f32o	3.65	0.0562	2.1634	1.0419	4.4919
f32u	0.02	0.8767	0.8245	0.3195	2.1275

difmh shows that f32b and f32j may show DIF according to age55 indicated by the p-values 0.0468 and 0.0449, respectively.

For f32b OR= 0.4283 with 95% CI (0.1988 ; 0.9225) - the estimate is outside the interval 0.65-1.53, but the CI shows possible values well above 0.65, so the magnitude may not be problematic. For item f32j OR=4.1044 with 95% CI (1.1434 ; 1 4.7325). Again we see the estimate is outside the interval 0.65-1.53, but the upper value 1.53 is included in the CI. Notice the opposing directions the ORs take. In one item the younger group has less probability of a positive answer than the older group. In the other item the younger group have larger probability of a positive answer than the older group.

Exercise 5
Investigate for DIF for age above and below 55 years using the logistic regression method. Are there any changes? Use the ORs to evaluate the magnitude.

```
* DIF investigated via logistic regression method
diflogistic f32a f32b f32c f32d f32f f32j f32l f32o f32u ///
    , group(age55)
```

```
Logistic Regression DIF Analysis
```

Item	Nonuniform Chi2	Nonuniform Prob.	Uniform Chi2	Uniform Prob.
f32a	0.08	0.7755	2.07	0.1502
f32b	1.13	0.2872	6.25	0.0124
f32c	0.60	0.4398	0.72	0.3952
f32d	0.04	0.8488	1.79	0.1811
f32f	3.80	0.0512	0.27	0.6041
f32j	1.78	0.1817	4.47	0.0346
f32l	0.05	0.8198	1.58	0.2086
f32o	2.77	0.0961	4.93	0.0264
f32u	0.56	0.4550	0.08	0.7746

diflogistic shows that none of the items display non-uniform DIF – none of the p-values in the leftmost column are below 0.05. The items f32b, f32j and f32o may display uniform DIF as indicated by p-values in the second column: 0.0124, 0.0346, and 0.0264 respectively. .

```
* In order to evaluate OR for uniform DIF
* for that we need the sumscore
egen sumscore_9dep=rowtotal(f32a f32b f32c f32d f32f f32j f32l f32o f32u)
logit f32b b1.age55 c.sumscore_9dep, or
logit f32j b1.age55 c.sumscore_9dep, or
logit f32o b1.age55 c.sumscore_9dep, or
```

The estimates of the ORs fall outside the interval 0.65-1.53, but again the CI are overlapping indicating that the magnitude may not be problematic.
F32b: OR= 2.61 95% CI (1.21 ; 5.62)
F32j: OR= 0.30 95% CI (0 .10 ; 0 .96)
F32o: OR= 0.44 95% CI (0..21 ; 0.93)

Exercise 6
Investigate the overall impact for one of the items by calculating the difference in pseudo R2.

```
* the model including non-uniform DIF
logit f32f b1.age55##c.sumscore_9dep
* the pseudo R2 is listed - see if you can find the estimate in the output
* from the logit command
display as text " R2 full dif= " as result e(r2_p)

* the model with no DIF
logit f32f c.sumscore_9dep
display as text " R2 no dif= " as result e(r2_p)

* here the difference is estimated
display  .43808584 - .43010923
 .00797661
```

The difference is 0.008 substantially below 0.03 indicating negligible DIF toward agg55 for item f32f.

For item f32b:

```
logit f32b b1.age55##c.sumscore_9dep
display as text " R2 full dif= " as result e(r2_p)
logit f32b c.sumscore_9dep
display as text " R2 no dif= " as result e(r2_p)
display .57830477 - .56239329
 .01591148
```

For item f32j:

```
logit f32j b1.age55##c.sumscore_9dep
display as text " R2 full dif= " as result e(r2_p)
logit f32j c.sumscore_9dep
display as text " R2 no dif= " as result e(r2_p)
display  .44132524 - .40513704
 .0361882
```

So item f32j seems to display uniform DIF with substantial impact, since the difference in pseudo R2 is above 0.03. Item f32j may be a candidate for item reduction, but again do not make such a decision based on statistics alone!

Exercise 7

Examine whether items f32a and f32b display signs of local dependence when using the partial correlation method. Do you come to similar conclusions when applying the logistic regression method?

```
* here scores  are estimated without item f32a and f32b, respectively
egen sumscore_8dep_f32a=rowtotal(f32b f32c f32d f32f f32j f32l f32o f32u)
egen sumscore_8dep_f32b=rowtotal(f32a f32c f32d f32f f32j f32l f32o f32u)

* here the partial correlation is estimated  - the correlation between f32a
* and f32b when taken the variable sumscore_8dep_f32a or sumscore_8dep_f32b
* into account.
pcorr f32a f32b sumscore_8dep_f32a
pcorr f32a f32b sumscore_8dep_f32b
```

```
. pcorr f32a f32b sumscore_8dep_f32a
(obs=482)

Partial and semipartial correlations of f32a with

                Partial   Semipartial      Partial   Semipartial   Significance
   Variable       Corr.         Corr.      Corr.^2       Corr.^2          Value
-------------------------------------------------------------------------------
        f32b      0.2666        0.2096       0.0711        0.0439         0.0000
  sumscor~32a     0.3259        0.2613       0.1062        0.0683         0.0000

. pcorr f32a f32b sumscore_8dep_f32b
(obs=482)

Partial and semipartial correlations of f32a with

                Partial   Semipartial      Partial   Semipartial   Significance
   Variable       Corr.         Corr.      Corr.^2       Corr.^2          Value
-------------------------------------------------------------------------------
        f32b      0.2354        0.1641       0.0554        0.0269         0.0000
  sumscor~32b     0.5341        0.4282       0.2852        0.1833         0.0000
```

Since both partial correlations are different from zero with both p-values<0.0001, f32a and f32b exhibits local dependence.

* via logit
logit f32a b1.f32b c.sumscore_8dep_f32b, or
logit f32b b1.f32a c.sumscore_8dep_f32a, or

Here both the estimates and CI are outside the interval 0.65-1.53 indicating uniform DIF with a substantial impact and thereby there are clear indication of local dependency between item f32a and f32b.
OR=0.29 95% CI (0.13 ; 0.63)
OR=0.18 95% CI (0.09 ; 0.39)