Problem Set #2
Statistics for Psychosocial Research

Q1: Using 10 items from 1996 self-esteem scale and 7 items from 1996 cesd depression scale:

Q1a: Perform principal components analysis on the 17 items and report results.

```
. factor slf961-slf9610 cesd961-cesd967, pc
(obs=328)

            (principal components; 17 components retained)
Component    Eigenvalue     Difference    Proportion    Cumulative
-------------------------------------------------------------------
     1        5.45697         3.38505      0.3210         0.3210
     2        2.07192         0.98650      0.1219         0.4429
     3        1.08543         0.14207      0.0638         0.5067
     4        0.94336         0.08347      0.0555         0.5622
     5        0.85989         0.06878      0.0506         0.6128
     6        0.79111         0.01412      0.0465         0.6593
     7        0.77699         0.04307      0.0457         0.7050
     8        0.73392         0.06046      0.0432         0.7482
     9        0.67346         0.04020      0.0396         0.7878
    10        0.63326         0.03677      0.0373         0.8251
    11        0.59649         0.03385      0.0351         0.8602
    12        0.56264         0.08853      0.0331         0.8933
    13        0.47411         0.08715      0.0279         0.9212
    14        0.38696         0.04075      0.0228         0.9439
    15        0.34621         0.01290      0.0204         0.9643
    16        0.33331         0.05935      0.0196         0.9839
    17        0.27396               .      0.0161         1.0000
```

Eigenvectors

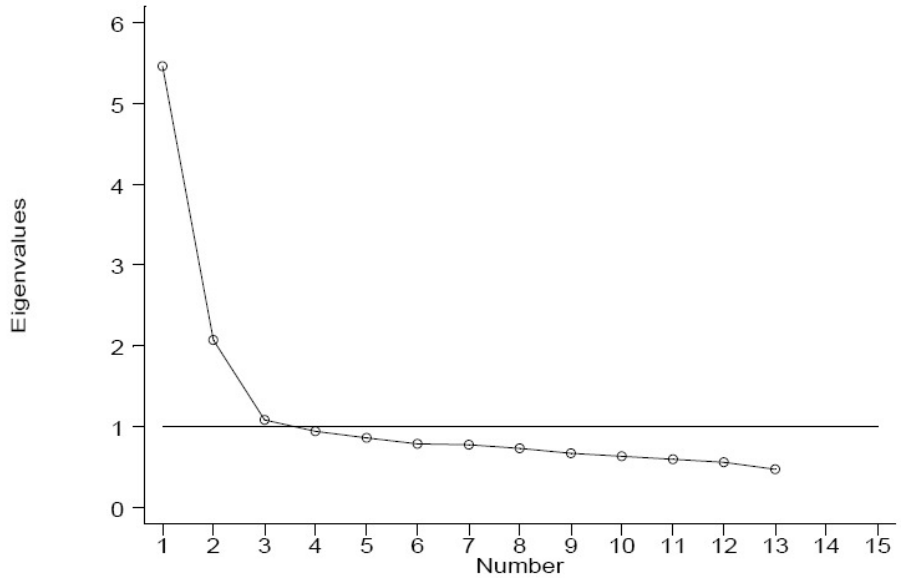
Variable	1	2	3	4	5	6
slf961	0.28044	0.14473	-0.16185	-0.03441	0.02357	0.01408
slf962	0.29673	0.22725	-0.14636	-0.14456	-0.13205	0.13413
slf963	0.29900	0.21821	0.11773	0.16446	-0.07803	-0.10378
slf964	0.24485	0.24869	-0.11704	-0.31829	-0.21900	-0.06208
slf965	0.29822	0.14071	0.05874	0.29646	0.07380	-0.04096
slf966	0.30615	0.16214	-0.15929	-0.18852	-0.00570	-0.05965
slf967	0.29163	0.15683	-0.17592	-0.17153	-0.03929	-0.03454
slf968	0.24185	-0.10054	-0.14358	0.45309	-0.07066	-0.03856
slf969	0.29495	-0.03822	0.26692	0.20726	0.14390	0.17236
slf9610	0.29598	0.04788	0.24053	0.16997	0.37226	0.19824
cesd961	-0.12805	0.26571	0.32528	0.41596	-0.58847	-0.00382
cesd962	-0.14509	0.35091	-0.04371	0.08780	0.28614	-0.61162
cesd963	-0.20996	0.40378	-0.21065	0.01634	-0.09543	-0.10387
cesd964	0.01033	0.21592	0.73269	-0.36951	0.09050	-0.18035
cesd965	-0.15982	0.31734	0.04166	-0.16364	0.09441	0.63902
cesd966	-0.20813	0.37374	-0.06859	0.15664	-0.10955	0.24228
cesd967	-0.18067	0.29937	-0.14133	0.22522	0.53877	0.07623

Variable	Eigenvectors 7	8	9	10	11	12
slf961	-0.33268	-0.13324	-0.44685	0.43366	0.31480	0.14253
slf962	-0.18839	-0.16452	-0.13863	0.23626	-0.13544	-0.01009
slf963	-0.09161	-0.27716	0.03747	-0.33003	-0.10122	-0.00017
slf964	0.17444	-0.00222	0.12351	0.00934	-0.00775	-0.76026
slf965	-0.15235	-0.36470	0.16176	-0.22427	-0.36590	0.05833
slf966	0.18003	0.17490	0.18393	-0.02083	-0.07155	0.38906
slf967	0.15984	0.43913	0.10412	0.00066	-0.07236	0.30544
slf968	0.24747	-0.15920	0.48398	0.21161	0.51953	0.03587
slf969	0.09487	0.37669	-0.19622	-0.10277	0.29671	-0.27140
slf9610	0.07241	0.23057	-0.23202	-0.17223	-0.10322	-0.00548
cesd961	-0.11920	0.33820	0.01242	0.29982	-0.22036	0.01610
cesd962	-0.42937	0.21207	0.08226	-0.12682	0.25242	-0.09271
cesd963	0.34563	0.07489	-0.12990	-0.24458	0.06233	0.11958
cesd964	0.23438	-0.24088	0.10188	0.21576	0.16344	0.16631
cesd965	-0.38776	0.08016	0.45252	-0.10837	0.18698	0.02926
cesd966	0.26136	-0.26682	-0.34778	-0.21811	0.27025	0.06622
cesd967	0.25832	-0.03033	0.08453	0.48960	-0.32946	-0.14998

Two factors explain 44.29% of the variance in the items; three factors explain 50.67% of the variance.

Q1b: Perform a scree plot and report the results.

See graph below. The scree plot shows two points above the "elbow" and three points above 1.



Q1c: How many factors do these items index? Why?

These items index two factors, because two factors have eigenvalues greater than 1 and are above the elbow on the scree plot.

(A case could also have been made for a third factor based on the eigenvalue rule alone, though, substantively, because two scales are in here, it really probably should be 2).

Q1d: If you were to redo the analysis with the number of factors you have selected, what method would you choose (e.g. maximum-likelihood, iterated principal likelihood, etc.)? What are the theoretical benefits of the method you chose over the others?

IPF would probably be best; the iterative nature may be more likely to arrive at the correct solution, and ML would require a stronger assumption of normality of the factors, and this has not been confirmed.

Q1e: Redo factor analysis, using method chosen in Q1d and constrain the solution to the number of factors chosen in Q1c.

```
. factor slf961-slf9610 cesd961-cesd967, ipf fa(2)
(obs=328)
```

Q1f: Rotate the factors and report the results, using either promax or varimax rotation; justify choice between these two methods.

```
. rotate
            (varimax rotation)
            Rotated Factor Loadings
 Variable |      1          2     Uniqueness
----------+-----------------------------------
   slf961 |   0.14450    0.62434    0.58931
   slf962 |   0.05970    0.72984    0.46377
   slf963 |   0.08080    0.72334    0.47025
   slf964 |  -0.00328    0.61034    0.62748
   slf965 |   0.17111    0.66368    0.53024
   slf966 |   0.14581    0.70374    0.48348
   slf967 |   0.13958    0.66153    0.54290
   slf968 |   0.34916    0.39994    0.71814
   slf969 |   0.37146    0.53954    0.57092
  slf9610 |   0.27711    0.59404    0.57033
  cesd961 |  -0.36456   -0.10498    0.85607
  cesd962 |  -0.48734   -0.08570    0.75515
  cesd963 |  -0.71592   -0.15030    0.46486
  cesd964 |  -0.15754    0.11045    0.96298
  cesd965 |  -0.46644   -0.13163    0.76511
  cesd966 |  -0.64755   -0.17368    0.55052
  cesd967 |  -0.48398   -0.17608    0.73476
```

I rotated using both varimax and promax rotations and compared the results; the different rotational methods did not change the loadings or uniquenesses, so in the interest of simplicity of interpretation, I will use varimax. I do not believe there is a problem with interdependence among the factors, which would necessitate use of the promax rotation.

The case for promax rotation could also have been made, either by mentioning substantively why the two constructs might be related, or by showing that the two total scale scores were related (if they were).

Q1g: Based on these results, should any items be removed from the analysis? If so, which one(s)?

Based on the above rotation, I would possibly remove slf968 and cesd964, because these are the items with the lowest loadings on each of the factors and have high uniquenesses.

Any items which could be shown to be redundant (v. highly correlated with another of the items) might also be removed.

Q1h: Using the codebook, look at the wording of the two removed questions and speculate on why these items are not consistent with the others.

Item slf968 deals with respect, which may be measuring the same thing as a number of the other questions, like slf966 and slf967 (positive attitude/satisfied with self). Item cesd964 (everything is an effort) could be measuring the same thing as item cesd967 (could not get going).

I think this might be wrong – I should think that if the items were highly correlated with others, they would show low uniqueness…

CESD4 4th question: I felt that everything I did was an effort.
SLF8 8th question (reversed): I wish I could have more respect for myself.

Q1i: Redo factor analysis, excluding items removed in (g).

```
. factor slf961-slf967 slf969 slf9610 cesd961-cesd963 cesd965-cesd967, ipf fa(2)
(obs=328)
```

[OUTPUT OMITTED]

```
. rotate

            (varimax rotation)
            Rotated Factor Loadings
 Variable |      1          2    Uniqueness
----------+--------------------------------
    slf961 |   0.14435    0.62776    0.58508
    slf962 |   0.06550    0.73633    0.45353
    slf963 |   0.09250    0.71449    0.48095
    slf964 |   0.00778    0.60999    0.62785
    slf965 |   0.16963    0.65627    0.54053
    slf966 |   0.14854    0.70338    0.48319
    slf967 |   0.14049    0.66611    0.53656
    slf969 |   0.37933    0.52918    0.57608
   slf9610 |   0.29550    0.58903    0.56572
   cesd961 |  -0.35652   -0.10896    0.86102
   cesd962 |  -0.47788   -0.08183    0.76493
   cesd963 |  -0.73750   -0.13931    0.43669
   cesd965 |  -0.44941   -0.12817    0.78160
   cesd966 |  -0.65411   -0.16841    0.54378
   cesd967 |  -0.49355   -0.17310    0.72645
```

There is not much impact on the other remaining variables in terms of loadings/uniqueness so can keep these two variables out.

Q1j: If necessary, repeat steps g-i until you arrive at a final pool of items.
Remove cesd961 as low loading and high uniqueness.

```
. factor slf961-slf967 slf969 slf9610 cesd962 cesd963 cesd965-cesd967, ipf fa(2)
(obs=328)
```

[OUTPUT OMITTED]

```
. rotate
            (varimax rotation)
            Rotated Factor Loadings
 Variable |      1          2    Uniqueness
----------+--------------------------------
    slf961 |   0.14459    0.62731    0.58557
    slf962 |   0.07064    0.73674    0.45222
    slf963 |   0.11388    0.70945    0.48371
    slf964 |   0.00713    0.61205    0.62535
    slf965 |   0.18530    0.65198    0.54059
    slf966 |   0.14162    0.70529    0.48251
    slf967 |   0.14174    0.66649    0.53570
    slf969 |   0.40310    0.52221    0.56481
   slf9610 |   0.30497    0.58446    0.56540
   cesd962 |  -0.47135   -0.07823    0.77171
   cesd963 |  -0.74478   -0.12964    0.42849
   cesd965 |  -0.44279   -0.12499    0.78831
   cesd966 |  -0.64280   -0.16432    0.55981
   cesd967 |  -0.50448   -0.16609    0.71791
```

Removing this additional variable also did not cause much change in the other variables, so seems okay to keep it out.

Final pool of items: self-esteem items 1-7, 9, 10 and depression items 2, 3, 5-7

Q2a: What type of validity would be involved if a researcher wanted to know how well depression items corresponded with a psychiatrist's diagnosis?

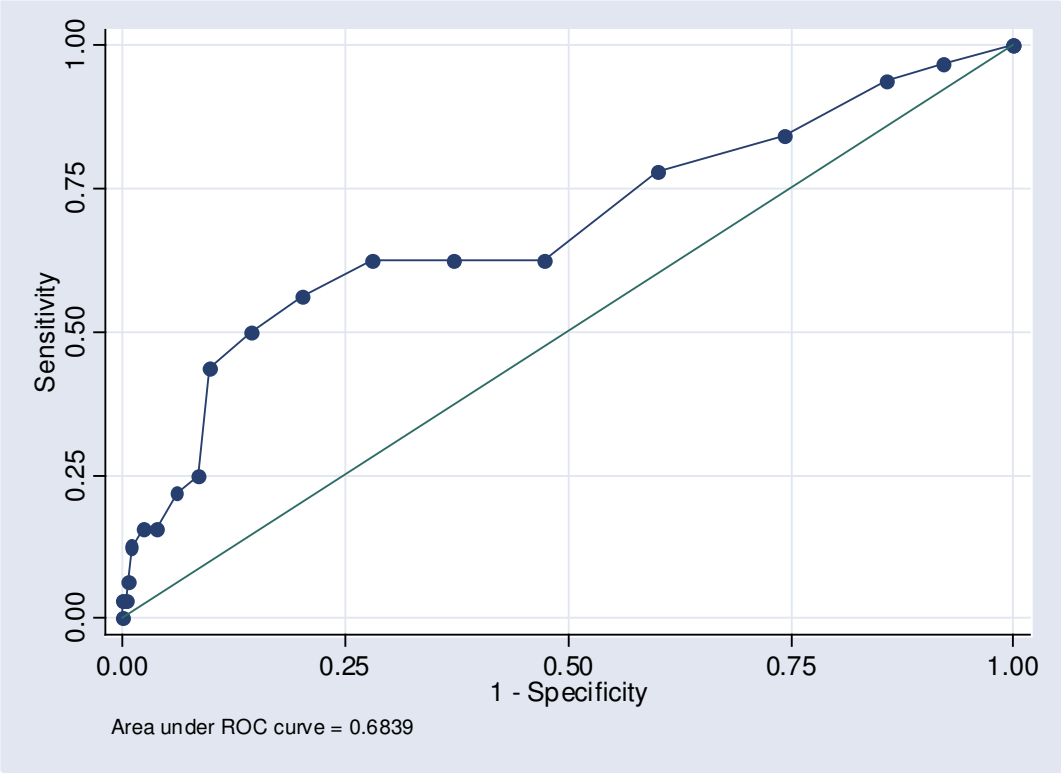
Criterion validity (concurrent) would be the type of validity involved, because the researcher would be comparing the depression items against a gold standard (the psychiatrist's diagnosis).

Q2b: What statistical procedure would help determine which cutoff on the1996 depression scale corresponds with a psychiatrist's diagnosis with minimum amount of error?

You could conduct the ROC curve procedure and then use the sensitivity/specificity calculated as part of the procedure to determine the appropriate cutoff point.

```
. gen cesdt =  cesd961+ cesd962+ cesd963+ cesd964+ cesd965+ cesd966+ cesd967

. logistic  psydp96 cesdt

Logistic regression                               Number of obs   =        328
                                                  LR chi2(1)      =      17.11

                                                  Prob > chi2     =     0.0000
Log likelihood = -96.303179                       Pseudo R2       =     0.0816
```

```
------------------------------------------------------------------------------
     psydp96 | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       cesdt |   1.205057   .0543898     4.13   0.000     1.103034    1.316516
------------------------------------------------------------------------------

. lroc

Logistic model for psydp96

number of observations =      328
area under ROC curve   =   0.6839
```



Detailed report of Sensitivity and Specificity

Cut point	Sensitivity	Specificity	Correctly Classified	LR+	LR-
(>= 0)	100.00%	0.00%	9.76%	1.0000	
(>= 1)	96.88%	7.77%	16.46%	1.0504	0.4022
(>= 2)	93.75%	14.19%	21.95%	1.0925	0.4405
(>= 3)	84.38%	25.68%	31.40%	1.1352	0.6086
(>= 4)	78.13%	39.86%	43.60%	1.2992	0.5487
(>= 5)	62.50%	52.70%	53.66%	1.3214	0.7115
(>= 6)	62.50%	62.84%	62.80%	1.6818	0.5968
(>= 7)	62.50%	71.96%	71.04%	2.2289	0.5211
(>= 8)	56.25%	79.73%	77.44%	2.7750	0.5487
(>= 9)	50.00%	85.47%	82.01%	3.4419	0.5850
(>= 10)	43.75%	90.20%	85.67%	4.4655	0.6236
(>= 11)	25.00%	91.55%	85.06%	2.9600	0.8192
(>= 12)	21.88%	93.92%	86.89%	3.5972	0.8318
(>= 13)	15.63%	96.28%	88.41%	4.2045	0.8763
(>= 14)	15.63%	97.64%	89.63%	6.6072	0.8642
(>= 15)	12.50%	98.99%	90.55%	12.3334	0.8840

```
( >= 16 )          6.25%        99.32%       90.24%       9.2500       0.9439
( >= 18 )          3.13%        99.66%       90.24%       9.2500       0.9720
( >= 20 )          3.13%       100.00%       90.55%                    0.9688
( >  20 )          0.00%       100.00%       90.24%                    1.0000
-------------------------------------------------------------------------------
```

Obs	ROC Area	Std. Err.	-Asymptotic Normal-- [95% Conf.	Interval]
328	0.6839	0.0569	0.57230	0.79552

Q2d: What cutoff score best corresponds with a psychiatrist's diagnosis with the least amount of error? How do you know?

One wants the point which corresponds is closest to the upper lefthand corner. In this case, it is

(>= 8) 56.25% 79.73% 77.44% 2.7750 0.5487.

It also has the highest sum of sensitivity plus specificity.

Q2e: What type of validity would be involved if a researcher wanted to develop a scale that indexed depression, and it was not possible to observe depression directly?

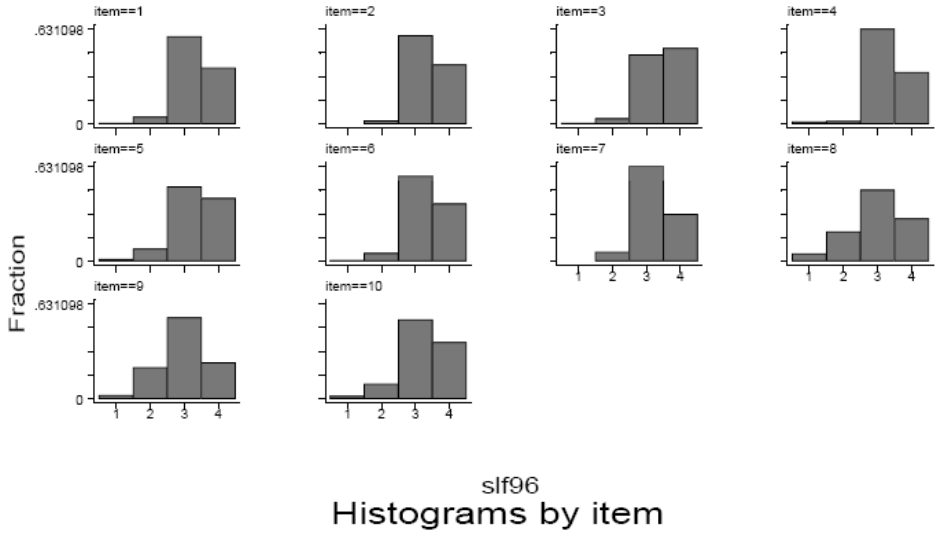
Construct validity, because you wouldn't have a gold standard with which to compare the scale.

Q2f: Given the answer to Q2e, the analyses in Q1 are most relevant to what specific types of validity, if one of the scales is intended to index depression?

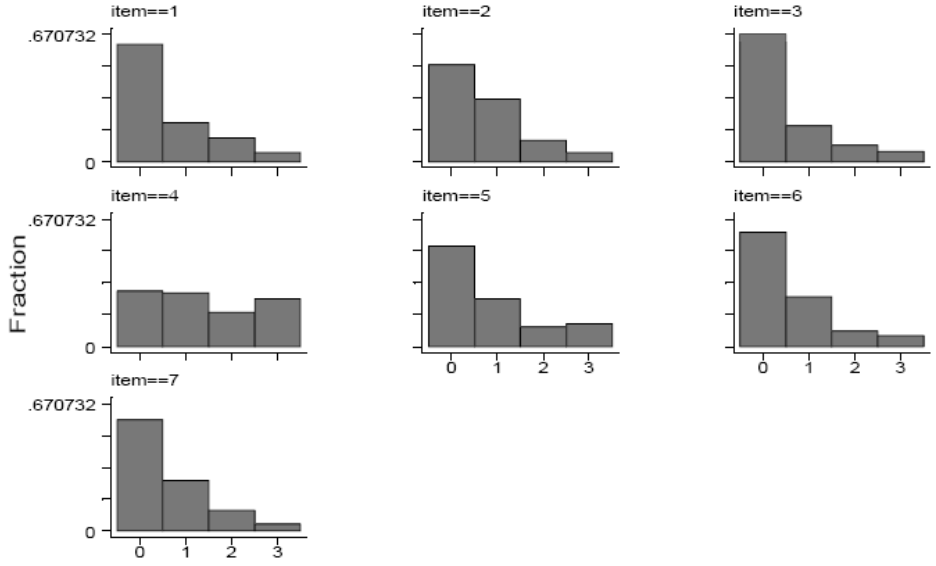
Internal construct validity, specifically convergent validity, because the factor analyses are trying to determine the extent the chosen items converge on the underlying depression trait.

You could also say discriminant construct validity, if you are talking about the differential loading of the two scales' items onto two different factors.

Q3a: Graph the histograms of the 1996 self-esteem items and report your results.



Q3b: Graph the histograms of the 1996 depression items and report your results.



Q3c: What assumptions of factor analysis might be violated by these variables? What are some possible consequences of these violations on the results?

The assumption of normality of the items (x's) may be violated. This might result in incorrect factor loading estimates.

Q3d: Print out the correlation matrix of the 1996 self-esteem and depression items and report your results.

{NOT SHOWN}

Q4a: In theory, how would you assess the external validity of the scales you created in Q1? Give specific examples of variables you might use if they were available.

You would want to assess how the scales corresponded with other known constructs and variables, eg. the relationship between depression and gender (would assume higher depression scores for females), and possibly compare to other validated depression and self-esteem measures currently in use. You could do this by administering both the known measure and your new measure to a sample population, and seeing if your new measures of depression and self-esteem corresponds to the results for each subject on the known measure.

Q4b: What would you do if your scales did not exhibit good external validity?

I would try again–by reevaluating the items in the scale against known measures and concepts behind depression and self-esteem and seeing if there were other items that should be considered and added to the survey for another round of data collection so that I could construct the scales again.

Q5a: Ran exploratory factor analysis in Mplus with continuous variable items.

[OUTPUT OMITTED]

Q5b: How does the correlation matrix compare to the results in Q3d?

The correlation matrix is basically the same in Mplus as in Stata, as is should be since the scale items are treated as continuous in both.

Q5c: How do the eigenvalues compare to your results in Q1a?

The eigenvalues are also the same in Mplus as in Stata.

Q5d: Ran exploratory factor analysis in Mplus with categorical variable items.

[OUTPUT OMITTED]

Q5e: How does the correlation matrix compare to your results in Q3d? Do you see any general trends?

The correlations are higher in Mplus than in Stata, by about .10 in each case.

Q5f: How do the eigenvalues compare to your results in Q1a? How do your conclusions change if you treat the items as categorical instead of continuous?

The eigenvalues are higher for the first two components (6.749 and 2.293 in Mplus vs. 5.456 and 2.071 in Stata) but about the same for the third and below. It does not appear that my conclusions change if I treat the items as categorical instead of continuous, but it does seem that the support for my conclusions is a little bit stronger–stronger correlations between the variables, and higher eigenvalues for the two factors model.