

一、選擇題 (40 分) 請選出最佳的一個選項

1. The \_\_\_\_\_ details tests in order to help the consumer select an appropriate test.
  - a) Measurement Handbook
  - b) APA Standards of Testing
  - c) Tests and Measurement Catalog
  - d) Mental Measurements Yearbook
  
2. Individuals with training versus without training in psychology are
  - a) more likely to recognize the limitations of tests they use
  - b) only qualified to use tests if they are licensed
  - c) less likely to use tests which require a great deal of interpretation
  - d) both b and c
  
3. The first scale to provide a practical and reasonably valid measure of intelligence was developed by
  - a) Stanford and Binet
  - b) Binet and Swallow
  - c) Galton and Cattell
  - d) Binet and Simon
  
4. Percentile transformations change the relationship between scores such that scores
  - a) near the center of the distribution differ substantially in their percentile ranks
  - b) near the center of the distribution are similar in their percentile ranks
  - c) at the extremes of the distribution differ substantially in their percentile ranks
  - d) throughout the same distribution have equivalent percentile ranks
  
5. The techniques used to equate scores across groups of examinees or forms of tests involve
  - a) comparing z-scores instead of raw scores
  - b) using anchor items
  - c) having all groups take all forms of the test
  - d) both a and b
  
6. A second grade child's score on a vocabulary test has a grade equivalent of 4.0. This child
  - a) should be in fourth grade
  - b) should be moved to third grade at least
  - c) is achieving at the overall level of most children entering fourth grade
  - d) earned a score that we would expect a typical fourth grader to earn on this test

7. In developing norms, it is important to
  - a) select samples that represent the target population
  - b) obtain a representative sample from the general population
  - c) obtain educational backgrounds on all samples
  - d) all of the above
  
8. In contrast to norm-referenced tests which assess one's performance relative to other individuals, criterion-referenced tests
  - a) assess one's performance in comparison to some standard(s)
  - b) assess one's performance in comparison to groups of people
  - c) evaluate one's performance within a group setting
  - d) are less challenging to develop
  
9. "True score" on a measure of anxiety represents
  - a) a person's real level of anxiety
  - b) a person's obtained level of anxiety
  - c) a combination of all factors leading to consistency in anxiety scores
  - d) both a and c
  
10. The standard error of measurement is a function of which two factors?
  - a) reliability of the test and range of test scores
  - b) variability of test scores and range of test scores
  - c) reliability of the test and variability of test scores
  - d) variability of test scores and sample size
  
11. A method of estimating the true correlation between X and Y given the correlation between two unreliable measures of X and Y is by using the
  - a) shrinkage coefficient
  - b) reliability coefficient
  - c) correction for attenuation
  - d) corrected reliability index
  
12. If individual differences are \_\_\_\_\_, reliability will be \_\_\_\_\_.
  - a) small; low
  - b) small; high
  - c) large; low
  - d) large; zero
  
13. Reliability of a test can be increased by
  - a) decreasing inter-item correlations
  - b) decreasing number of test items
  - c) increasing number of test items
  - d) both a and c

14. What is "validated" in a validity study?
  - a) the test
  - b) test scores
  - c) the inferences reached on the basis of test scores
  - d) all of the above
  
15. The most appropriate method(s) for construct validation is(are) the
  - a) correlation coefficient
  - b) laboratory experiment
  - c) multitrait-multimethod approach
  - d) all are appropriate methods
  
16. The ideal criterion-related validity strategy is
  - a) predictive validity
  - b) concurrent validity
  - c) construct validity
  - d) face validity
  
17. Range restriction is most likely to occur in
  - a) concurrent validity studies
  - b) predictive validity studies
  - c) construct validity studies
  - d) both a and b
  
18. Item Response Theory provides measures that are
  - a) sample specific
  - b) sample indifferent
  - c) sample invariant
  - d) sample inversed
  
19. When the item-total correlation is \_\_\_\_\_, the slope of the item characteristic curve is \_\_\_\_\_.
  - a) negative; positive
  - b) negative; negative
  - c) positive; negative
  - d) positive; zero
  
20. Computerized adaptive tests (CAT) attempt to
  - a) measure an individual's ability to react to various stimuli
  - b) tailor test items to an individual's ability level
  - c) obtain a wide variety of information about an individual in a relatively short period of time
  - d) allow individuals to choose a test format

## 二、問答題

Suppose a two-parameter logistic model is fit to data for a test. For the first two items  $P_1(\theta) = 0.5$  and  $P_2(\theta) = 0.7$  for a particular latent trait score,  $\theta$ . In addition, for these two items,  $P(+, +|\theta) = 0.45$  and  $P(-, +|\theta) = 0.05$ , where '+' is denoted as a correct answer and '-' as an incorrect answer. Do these results suggest the test is unidimensional? Why or why not? (10 分)

三、需要進行統計檢定時請設定  $\alpha = 0.05$ ，並使用最後的附表

- 想像你到了小人國，隨機抽取 8 位小人國的人，得到身高為 9.0, 7.6, 8.4, 12.0, 6.4, 9.6, 4.2 和 6.8 公分。試考驗小人國的人平均身高不及 10 公分這種說法是否成立？〔10 分〕
- 某高中老師調查 100 位高中學生有無打工經驗，得到下表結果。試考驗高一、高二以及高三學生有過打工經驗的比率是否相同？〔10 分〕

打工經驗	高一	高二	高三	人數合計
有	10	15	20	45
無	30	15	10	55
年級人數合計	40	30	30	

- 消費者保護機構檢驗市售汽車電池的使用期限，隨機抽樣 A、B、C 三家廠商各 10 顆電池，並觀察其使用壽命，得到數據如下所示。請使用變異數分析判斷三廠牌電池使用壽命是否有差異。〔15 分〕

廠牌	A	B	C
樣本數	10	10	10
平均壽命	16	20	24
樣本變異數	20	25	30

- 某年學測英文作文有兩位老師共同批改。為了解兩位老師評分標準是否一致，先隨機抽取 10 份考卷由兩位老師共同批改之後，得到下表分數。請問依據這個樣本，是否能斷定兩位閱卷老師的評分標準不一致？〔15 分〕

學生	一	二	三	四	五	六	七	八	九	十
老師 A	5	10	9	12	13	14	9	12	9	8
老師 B	6	12	10	13	8	10	6	8	12	6

國立中正大學九十八學年度碩士班招生考試試題

系所別：心理學系臨床心理學

科目：心理統計學與心理測驗學

第3節

第5頁，共5頁

Critical Values for t

	單尾 $\alpha$	0.05	0.025	0.01	0.005
df	雙尾 $\alpha$	0.10	0.05	0.025	0.01
2		2.92	4.3027	6.2054	9.925
3		2.3534	3.1824	4.1765	5.8408
4		2.1318	2.7765	3.4954	4.6041
5		2.015	2.5706	3.1634	4.0321
6		1.9432	2.4469	2.9687	3.7074
7		1.8946	2.3646	2.8412	3.4995
8		1.8595	2.306	2.7515	3.3554
9		1.8331	2.2622	2.685	3.2498
10		1.8125	2.2281	2.6338	3.1693

Critical Values for  $\chi^2$

df	$\alpha$	0.1	0.05	0.025	0.01	0.005
1		2.70554	3.84146	5.02389	6.6349	7.87944
2		4.60517	5.99146	7.37776	9.21034	10.59663
3		6.25139	7.81473	9.3484	11.34487	12.83816
4		7.77944	9.48773	11.14329	13.2767	14.86026
5		9.23636	11.0705	12.8325	15.08627	16.7496

F table for  $\alpha = 0.05$

df2/df1	1	2	3	4	5	6
26	4.2252	3.369	2.9752	2.7426	2.5868	2.4741
27	4.21	3.3541	2.9604	2.7278	2.5719	2.4591
28	4.196	3.3404	2.9467	2.7141	2.5581	2.4453
29	4.183	3.3277	2.934	2.7014	2.5454	2.4324
30	4.1709	3.3158	2.9223	2.6896	2.5336	2.4205