

101 年公務人員普通考試試題

代號：41060

全二頁
第一頁

類 科：教育行政

科 目：教育測驗與統計概要

考試時間：1 小時 30 分

座號：

※注意：(一)可以使用電子計算器。

(二)不必抄題，作答時請將試題題號及答案依照順序寫在試卷上，於本試題上作答者，不予計分。

一、請闡述實作評量的特性。(25 分)

二、收集構念效度證據常用的方法有那幾種？試以「數學推理評量」為例，簡要說明之。(25 分)

三、已知一變數 X 成偏態分配，其平均數為 20，標準差為 5。(附常態分配表)

(一)從此分配抽取 8 個分數並計算其平均數，若重複此步驟無限多次，則這些樣本平均數所形成的分配之平均數為何？(5 分)

(二)從此分配只抽取 5 個分數並計算其平均數，再重複此步驟無限多次之後，這些樣本平均數所形成的分配之變異誤為何？(5 分)

(三)若從同一個分配抽取 100 個分數並計算其平均數，重複此步驟無限多次之後，可以得到一個由樣本平均數所形成的分配。在這個分配中，樣本平均數小於 20.98 的概率為何？(10 分)

四、下面左表為學生在數學推理及語文推理上的成績，右表為數學推理的部分成績。請根據數據回答問題，列出算式並解釋結果(未列出算式者，不予計分)。(附常態分配表)

			數學推理部分成績		
	數學推理	語文推理	原始分數	次數	累積次數
人數	100	100	...		
平均數	48	60	55	3	80
標準差	8	10	54	4	77
中數	49	58	53	2	74
眾數	50	56	...		

(一) A 生的數學推理與語文推理成績同為 54 分，A 生在那一個測驗上的表現較好？(5 分)

(二)承上題，A 生數學推理的 PR 值是多少？請解釋其表現。(5 分)

(三)若 B、C、D、E 四人在數學推理測驗上的 PR 值分別為 10、20、50、60，則 B、C 原始分數的差異是否與 D、E 原始分數的差異相同？(5 分)

(四)教師將數學推理測驗的原始分數做常態化轉換，形成 L 量尺分數，設定其分配之平均數為 80、標準差為 10。A 生的數學推理 L 量尺分數是多少？(5 分)

(五)數學推理測驗分數最低 10% 的學生需要補救教學。教師該設定補救教學的 L 量尺截點分數是多少？(5 分)

(六)數學資優生在數學推理測驗上的最低標準是 L 量尺分數 100 分。全校 1000 名學生中有多少位是資優生？(5 分)

(請接第二頁)

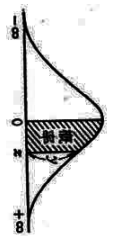
101年公務人員普通考試試題

代號：41060

全二頁
第二頁

類科：教育行政
科目：教育測驗與統計概要

z	標準	y	z	標準	y	z	標準	y	z	標準	y	z	標準	y	z	標準	y
00	0.000	3989	50	1.000	1915	100	2.000	4772	150	3.000	6987	200	4.000	8997	250	5.000	10000
01	0.004	3998	51	1.004	1921	101	2.004	4777	151	3.004	7000	201	4.004	9000	251	5.004	10004
02	0.008	3999	52	1.008	1927	102	2.008	4782	152	3.008	7013	202	4.008	9004	252	5.008	10008
03	0.012	3998	53	1.012	1933	103	2.012	4787	153	3.012	7026	203	4.012	9008	253	5.012	10012
04	0.016	3996	54	1.016	1939	104	2.016	4792	154	3.016	7039	204	4.016	9012	254	5.016	10016
05	0.019	3994	55	1.019	1945	105	2.019	4797	155	3.019	7052	205	4.019	9016	255	5.019	10019
06	0.023	3992	56	1.023	1951	106	2.023	4802	156	3.023	7065	206	4.023	9020	256	5.023	10023
07	0.027	3990	57	1.027	1957	107	2.027	4807	157	3.027	7078	207	4.027	9024	257	5.027	10027
08	0.031	3988	58	1.031	1963	108	2.031	4812	158	3.031	7091	208	4.031	9028	258	5.031	10031
09	0.035	3986	59	1.035	1969	109	2.035	4817	159	3.035	7104	209	4.035	9032	259	5.035	10035
10	0.038	3984	60	1.038	1975	110	2.038	4822	160	3.038	7117	210	4.038	9036	260	5.038	10038
11	0.042	3982	61	1.042	1981	111	2.042	4827	161	3.042	7130	211	4.042	9040	261	5.042	10042
12	0.046	3980	62	1.046	1987	112	2.046	4832	162	3.046	7143	212	4.046	9044	262	5.046	10046
13	0.051	3978	63	1.051	1993	113	2.051	4837	163	3.051	7156	213	4.051	9048	263	5.051	10051
14	0.057	3976	64	1.057	2000	114	2.057	4842	164	3.057	7169	214	4.057	9052	264	5.057	10057
15	0.056	3975	65	1.056	2006	115	2.056	4847	165	3.056	7182	215	4.056	9056	265	5.056	10060
16	0.056	3974	66	1.056	2012	116	2.056	4852	166	3.056	7195	216	4.056	9060	266	5.056	10063
17	0.055	3973	67	1.055	2018	117	2.055	4857	167	3.055	7208	217	4.055	9064	267	5.055	10066
18	0.054	3972	68	1.054	2024	118	2.054	4862	168	3.054	7221	218	4.054	9068	268	5.054	10069
19	0.053	3971	69	1.053	2030	119	2.053	4867	169	3.053	7234	219	4.053	9072	269	5.053	10072
20	0.053	3970	70	1.053	2036	120	2.053	4872	170	3.053	7247	220	4.053	9076	270	5.053	10075
21	0.052	3969	71	1.052	2042	121	2.052	4877	171	3.052	7260	221	4.052	9080	271	5.052	10078
22	0.052	3968	72	1.052	2048	122	2.052	4882	172	3.052	7273	222	4.052	9084	272	5.052	10081
23	0.051	3967	73	1.051	2054	123	2.051	4887	173	3.051	7286	223	4.051	9088	273	5.051	10084
24	0.048	3966	74	1.048	2060	124	2.048	4892	174	3.048	7299	224	4.048	9092	274	5.048	10087
25	0.048	3965	75	1.048	2066	125	2.048	4897	175	3.048	7312	225	4.048	9096	275	5.048	10090
26	0.047	3964	76	1.047	2072	126	2.047	4902	176	3.047	7325	226	4.047	9100	276	5.047	10093
27	0.046	3963	77	1.046	2078	127	2.046	4907	177	3.046	7338	227	4.046	9104	277	5.046	10096
28	0.045	3962	78	1.045	2084	128	2.045	4912	178	3.045	7351	228	4.045	9108	278	5.045	10099
29	0.044	3961	79	1.044	2090	129	2.044	4917	179	3.044	7364	229	4.044	9112	279	5.044	10102
30	0.044	3960	80	1.044	2096	130	2.044	4922	180	3.044	7377	230	4.044	9116	280	5.044	10105
31	0.043	3959	81	1.043	2102	131	2.043	4927	181	3.043	7390	231	4.043	9120	281	5.043	10108
32	0.043	3958	82	1.043	2108	132	2.043	4932	182	3.043	7403	232	4.043	9124	282	5.043	10111
33	0.042	3957	83	1.042	2114	133	2.042	4937	183	3.042	7416	233	4.042	9128	283	5.042	10114
34	0.041	3956	84	1.041	2120	134	2.041	4942	184	3.041	7429	234	4.041	9132	284	5.041	10117
35	0.041	3955	85	1.041	2126	135	2.041	4947	185	3.041	7442	235	4.041	9136	285	5.041	10120
36	0.040	3954	86	1.040	2132	136	2.040	4952	186	3.040	7455	236	4.040	9140	286	5.040	10123
37	0.040	3953	87	1.040	2138	137	2.040	4957	187	3.040	7468	237	4.040	9144	287	5.040	10126
38	0.039	3952	88	1.039	2144	138	2.039	4962	188	3.039	7481	238	4.039	9148	288	5.039	10129
39	0.039	3951	89	1.039	2150	139	2.039	4967	189	3.039	7494	239	4.039	9152	289	5.039	10132
40	0.038	3950	90	1.038	2156	140	2.038	4972	190	3.038	7507	240	4.038	9156	290	5.038	10135
41	0.038	3949	91	1.038	2162	141	2.038	4977	191	3.038	7520	241	4.038	9160	291	5.038	10138
42	0.037	3948	92	1.037	2168	142	2.037	4982	192	3.037	7533	242	4.037	9164	292	5.037	10141
43	0.037	3947	93	1.037	2174	143	2.037	4987	193	3.037	7546	243	4.037	9168	293	5.037	10144
44	0.036	3946	94	1.036	2180	144	2.036	4992	194	3.036	7559	244	4.036	9172	294	5.036	10147
45	0.036	3945	95	1.036	2186	145	2.036	4997	195	3.036	7572	245	4.036	9176	295	5.036	10150
46	0.035	3944	96	1.035	2192	146	2.035	5002	196	3.035	7585	246	4.035	9180	296	5.035	10153
47	0.035	3943	97	1.035	2198	147	2.035	5007	197	3.035	7598	247	4.035	9184	297	5.035	10156
48	0.034	3942	98	1.034	2204	148	2.034	5012	198	3.034	7611	248	4.034	9188	298	5.034	10159
49	0.034	3941	99	1.034	2210	149	2.034	5017	199	3.034	7624	249	4.034	9192	299	5.034	10162
50	0.033	3940	100	1.033	2216	150	2.033	5022	200	3.033	7637	250	4.033	9196	300	5.033	10165



標準分配表

(續)

(續)

□ 申論題解答

一、【擬答】

(一)實作評量的定義：

實作評量 (performance assessment) 又稱非紙筆測驗，係指根據學生實際完成一項特定任務或工作表現所作的評量。這些任務或工作，可能是實際操作、科學實驗、數學解題、口頭報告、寫作…等等。因此，其所使用的方式，係透過直接的觀察學生表現或間接的從學生作品去評量。這種評量方式不同於傳統的紙筆測驗，它重視實作的能力，就像駕照考試的路考一樣，係從實際的行為表現來評量，而不是依賴筆試來決定。

(二)實作評量的特性：

實作評量重視知識與能力的整體性，異於傳統紙筆測驗將知識分成片斷來練習。實作評量具有以下幾個特性：

1.實作的表現：

學生以實作的方式來表現他們對知識的理解，而實作評量的題目並不強調複雜性。

2.真實的情境：

實作評量講求日常生活的運用，讓問題建立在真實情境的脈絡上，使實作評量更具意義，協助學生更加理解與掌握問題的性質。

3.弱結構：

為了模擬日常生活中問題模糊不清的特性，實作評量的情境應傾向模糊，問題的設計傾向較不具結構的開放性問題，以測驗學生問題解決的創造性與統合性。

4.重視問題解決的過程與結果：

專家與生手的差別在於專家知道如何有效應用所習得的知識。實作評量能從學生操作的過程去瞭解學生的思考歷程，也能瞭解學生的操作是否符合程序。

5.重視小組的互動：

在真實的社會情境中，許多問題的解決必須透過社會互動的方式來達成，有些類型的實作工作經由小組互動，能激發彼此的創造力與問題解決能力，亦能培養學生人際互動的技巧。

6.時間彈性：

在極短的時間內評量學生的學習成效，往往過度窄化了學習的結果。通常高品質的答案需要相當的時間思考建構，應該儘量給予充裕的時間來解答，實作評量可以依照情境的需求，等待學生練習充裕之後再進行測試。

7.多向度的評分：

實作評量的內容多元且複雜，所以評量的標準相對複雜，而實作評量的評分至少應包含知識的產生、專業的研究及附帶價值三個層面。此外，實作評量結合真實生活的應用，教師將有意義的教學活動作為實作作業，要求學生表現、實作、產出或建構，使用人為判斷的方式計分，使得教師在教學與評量上均扮演與以往不同的角色。

二、【擬答】

(一)構念效度定義：

構念效度(或稱建構效度)是指測驗的編製過程中，以客觀方法驗證一份測驗是否符合理論假設，以反映該測驗

在建構結構上的有效程度。構念是由理論上所創造出來以理解人類行為或心理特質的概念，如果效度良好，則測驗所產生的分數即可真實反映受測者的特質，測驗者可依該構念之理論對所測得的結果加以詮釋。

(二)以「數學推理評量」為例，構念效度的驗證方法有下列五種。

1.內部凝聚性檢驗：

此類方法旨在於檢驗測驗內部之品質(如試題、分測驗、反應過程)是否測量同樣的心理構念。使用此法時以測驗總分為效標並計算各試題或分測驗與總分之關係，所得之相關係數即代表測驗內部之凝聚程度，凝聚程度越高即表示越傾向於測量同一心理構念。

2.外在相關係數檢驗：

驗證對象測驗與其他測驗工具間之關係。選用一已具有效度之測驗來做為效標，然後計算對象測驗與此效標測驗間之相關係數，即得測驗的效度，此一係數又稱符合效度。除此以外，相關法有包括以對比樣本為效標之效度檢驗，在使用此法時其主旨是衡量測驗對不同樣本之辨別能力。例如一組為數學推理能力高於平均值者，另一組為數學推理能力低於平均值者，若要檢驗某一數學推理量表之構念效度，如該量表能有效區別出兩樣本，即證明其能準確地測量心理特質。

3.因素分析：

為驗證以構念為中心的效度最常用的方法，它是一種多變數統計方法，可同時分析大量變項間之關係並找出他們在測量上的共同性。如果測驗原來產生的分數結構符合因素分析所產生的特質結構，則表示該測驗具有良好的因素效度。

4.幅合效度與辨別效度檢驗：

前者表示一測驗與其理論上同類之變項具有高相關，如數學推理測驗分數與在校數學科目成績有高相關，即為幅合效度。後者表示它與其在理論上並不相干之變項應不具相關，如數學推理測驗分數與在校閱讀理解測驗成績有低相關甚至零相關，即為辨別效度。可由「多元特質—多重方法矩陣」來同時檢驗這兩種效度。

5.發展現象與實驗效果之檢驗：

在心理學研究中有少數特質其發展過程與成長年齡或階段具有密切關係(例如智力、成就)。因此在檢查這類測驗的效度時，便可以比較不同年齡階層的受試樣本在測驗上的表現是否呈現如預期之顯著差異，如果結果確實如此，則測驗之效度即獲得肯定。最後，與構念有關的效度亦可藉著分析實驗變項對測驗分數影響之方法加以衡鑑。如果用來測驗此一變項之測驗具有效度，則其分數應能正確地反應研究假設中所推論的效果。

三、【擬答】

(一)中央極限定理，也稱為常態收斂定理，主要是指從平均數為 μ ，標準差為 σ 的母體中，隨機地抽取大小為 n 的獨立樣本 X_1, \dots, X_n 。當樣本數 n 很大時，其樣本平均減掉平均數，再除以標準差 σ/\sqrt{n} ，將會趨近平均數為 0，標準差為 1 的常態分佈(normal distribution)。題目僅取 8 個分數屬於小樣本，故母體不為常態分配，故無法計算。若取樣數 n 很大時，則樣本平均數為 $E(\bar{X})=20$ 。

(二)同(一)理由，僅取 5 個分數，母體不為常態分配，故無法計算。若取樣數 n 很大時，則樣本標準差為 $\sigma^2(\bar{X})=\frac{\sigma^2}{n}$

$$=\frac{5^2}{5}=5。$$

(三)依題意，查表可知(資料標準化)

$$P(\bar{X} \leq 20.98) = P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq \frac{20.98 - 20}{5/\sqrt{100}}\right) = P(\bar{Z} \leq 1.96) = 0.9756$$

四、【擬答】

(一) A 生分別對其推理測驗資料標準化

$$1. \text{數學推理測驗 } Z = \frac{X - \bar{X}}{S} = \frac{54 - 48}{8} = 0.75$$

$$2. \text{語文推理測驗 } Z = \frac{X - \bar{X}}{S} = \frac{54 - 60}{10} = -0.6$$

由於語文推理測驗之 Z 值較小(為負值)，故數學推理測驗表現較好。

(二)由(一)可知，數學推理測驗之 $Z=0.75$ ，經查表可得機率为 0.2734

則 PR 值为 $P(x \leq 54) = P(z \leq 0.75) = 0.5 + 0.2734 = 0.7734 = 77.34\%$ ，約為 77。

(三)不相同；因為資料標準化，故轉換單位不同，靠近常態分布中央原始分數單位較小，兩端點單位較大。

(四)依題意可得

原始分數	次數	累積次數	cf 中點	cp 中點
55	3	80		
54	4	77	75.5	0.755
53	2	74		

已知 cp 中點 = 0.755 經查表得 $Z=0.69$

且設定之平均數為 80，標準差為 10

則常態化量尺分數 = $10Z + 80 = 10 \times 0.69 + 80 = 86.9$ ，取為 87。

(五)數學推理測驗分數最低 10%，經查表 $Z = -1.28$ (對應機率为 0.1003)

則補救教學之量尺截點分數 = $10Z + 80 = 10 \times (-1.28) + 80 = 67.2$ ，取為 67。

(六)依題意知， $10Z + 80 = 100$ ，可得 $Z = 2$

則經查表 $P(Z > 2) = 0.0228$ ，故全校 1000 名學生中之資優生數目 = $1000 \times 0.0228 = 22.8$ ，取 22 人(注意，第 23 人錄取基準較 $Z=2$ 低，故取 22 人)。