題號： 345
科目： 統計學(H)
節次： 2

國立臺灣大學 111 學年度碩士班招生考試試題

題號：345

共 9 頁之第 1 頁

※ 注意：請於試卷內之「非選擇題作答區」標明題號依序作答。

(A) Short answer questions: （可參考題目最後的常態分配表）

1. (10%)
A service design researcher is interested in investigating the relationship between alcohol consumption and sleep disorders to develop an innovative app for potential users.
(a) Briefly describe how the researcher might design an observational study to investigate the relationship between alcohol consumption and sleep disorders?
(b) How can you show that consuming alcohol causes sleep disorders without being able to do a designed experiment?

2. (20%)
Many supermarkets offer customers discounts in return for using a shopper's card. These stores also have scanners equipment that records the item purchased by the customers. The combination of a shopping card and canner software allow the supermarket to track which items customers regularly purchase.

The data in the table below care based on items purchased in 86,214 shopping trips made by customers participating in a test of scanners. As part of the study, the customers also reported few things about themselves, such as the number of children and pets. The table shows the number of cats owned by the customer who applied for the shopping card. The table also shows in rows the number od cat food items purchases at the supermarket during each shopping trip in the time period of the study. Some particular type of probability is shown for each combination.

	No cats	1 cat	2 cats	3 cats	More than 3 cats
No cat food items	0.0487	0.0217	0.0025	0.0002	0
1 to 3	0.1698	0.0734	0.0104	0.0004	0.0002
4 to 6	0.1182	0.0516	0.0093	0.0006	0.0002
7 to 12	0.1160	0.0469	0.0113	0.0012	0.0005
More than 12	0.2103	0.0818	0.0216	0.0021	0.0011

(a) What type of probability is shown in this table?
(b) What could it mean that the probability of customers with no cats buying cat food is larger than zero?
(c) The smallest probabilities in the table are in the right-most column. Does this mean that owners of more than three cats buy relatively little cat food?
(d) Find the conditional probability of buying more than three cat food items among customers reported to own no cats. Compare this to the conditional probability of buying more than three cat food items among customers reported to own more than three cats.

見 背 面

題號： 345
科目： 統計學(H)
節次： 2

國立臺灣大學 111 學年度碩士班招生考試試題

題號：345

共 9 頁之第 2 頁

(e) What do you conclude about the use of this version of the scanner data to identify customers likely to purchase cat food?

3. (20%)

A computer peripheral manufacturer stockpiles monitors and printers in a large warehouse for shipment to retail stores. Some peripherals get damaged in handling. The long-term goal has been to keep the level of damaged machines below 2%. In a recent test, an inspector randomly checked 4 dozens of monitors and discovered that 4 of them had scratches or dents. Test the null hypothesis H_0: $p \leq 0.02$ in which p represents the probability of a damaged monitor.

(a) Do these data supply enough evidence to reject H_0? Use a binomial model to obtain the p-value.
(b) What assumption is necessary in order to use the binomial model for the count of the number of damaged washers?
(c) Test H_0 by using a normal model for the sampling distribution of $\hat{p}$. Does this test reject H_0 ?
(d) Which test procedure should be used to test H_0? Explain your choice.

(B) Multiple-Choice Questions (50%)
   (25 questions, two points each)

Please read the following articles and then answer the subsequent questions.

Just stepping into his office on the 68th floor of Willis Tower, Chicago in the early morning of Friday, January 28, 2022, Dave Toupet, co-founder and COO of MobiSmart, noticed the scene just right outside of the French window that the banks of mist drifted from Lake Michigan. As the view of the lakeshore got blurred, so was Toupet's concern on the next strategic move of MobiSmart, a startup offering the total solution of Advanced Driver Assistance Systems (ADAS). In a few hours, Toupet was going to brief the board of directors MobiSmart 2025, a business plan for MobiSmart to achieve leadership in the industry of autonomous vehicle (AV) by 2025. Due to potential astronomical investment and the high risk of failure, there had been serious debates among the top management of MobiSmart regarding which market and product should be the focus. MobiSmart might target on consumer AVs, long-haul trucks, or both. Each strategic option had its specific capital terms, operational challenges, risk factors, and profit prospects. Toupet's team had prepared three versions of business plans, but Toupet was still uncertain which one to go and whether the board would support any of his decisions.

Toupet was troubled by the contradictory 2025 forecasts made by MobiSmart's marketing department and a contracted marketing research firm. With differential parameters and

題號： 345
科目： 統計學(H)
節次： 2

國立臺灣大學 111 學年度碩士班招生考試試題

題號：345

共 9 頁之第 3 頁

assumptions on competitors' moves (CM), market potential in Different Segment (MPDS), horizontal and vertical partners' support (PS), and potential cannibalization of internal resources (PCIR), MobiSmart's marketing department predicted that the highest revenue would be achieved by focusing on consumer AVs that was expected to contribute US$3.4/2.8 billion of sales (in the promising/average condition) followed by long-haul truck with US$2.4/2.0 billion (in the promising/average condition) and both with US$ 2.0/1.8 billion (in the promising/average condition). However, the contracted marketing firm estimated MobiSmart's highest revenue of US$3.8/3.4 billion (in the promising/average condition) would be from the focus on long-haul truck, and the focus on both would lead to an expected revenue of US$2.6/2.2 billion (in the promising/average condition). The least expected revenue of US$ 1.2/.6 billion (in the promising/average condition) would be from the focus on consumer AVs. Toupet knew very well that MobiSmart's true market performance in 2025 might not be even close to either forecast.

1. Which of the followings can best describe the uncertainty for Toupet and his team to explore the unknown market responses on MobiSmart's offering in the process of hypothesis testing.
   A. "上善若水。水善利萬物而不爭，處眾人之所惡。" 老子，第八章
   B. "大盈若沖，大巧若拙。" 老子，第四十五章
   C. "天之蒼蒼，其正色邪？" 莊子，逍遙遊
   D. "以指喻指之非指，不若以非指喻指之非指也。" 莊子，齊物論

2. If Toupet would like to test the hypothesis of MobiSmart's expected revenue by focusing on consumer AVs, the power of the test would be increased by ____.
   A. collecting more sample participants, who are general consumers and are likely to adopt AVs by 2025, to examine the hypothesis
   B. speeding up the commercialization of the most advanced ADAS in the global market by 2025
   C. recruiting more consumers who participate in the online survey to examine the hypothesis
   D. identifying another third party that forecasted MobiSmart's highest revenue from the focus on consumer AVs by 2025

3. If Toupet worried much more about the potential loss of not being able to focus on consumer AVs, the significance level would be ____ when there was another reliable third party making the forecasts of MobiSmart's highest revenue from the focus on consumer AVs by 2025.
   A. Increased
   B. reduced
   C. unchanged
   D. equal to 0.1

見背面

題號： 345
科目： 統計學(H)
節次： 2

國立臺灣大學 111 學年度碩士班招生考試試題

題號：345

共 9 頁之第 4 頁

4. If Toupet's team found out that consumers AVs received significantly more attention than long-haul trucks in the Google Analytics, then _____.
   A. consumer AVs would generate much better market performance than long-haul trucks
   B. MobiSmart would be better off to focus on consumer AVs
   C. Toupet would be more confident to persuade MobiSmart's board of directors to invest in consumer AVs
   D. None of the above

5. When Toupet makes a hypothesis testing and concludes that _____, Type I error occurs.
   A. MobiSmart's revenue by focusing on consumer AVs is greater than that by focusing on long-haul trucks but it is in fact not greater.
   B. MobiSmart's revenue by focusing on consumer AVs is greater than that by focusing on long-haul trucks and it is in fact greater.
   C. MobiSmart's revenue by focusing on consumer AVs is not greater than that by focusing on long-haul trucks but it is in fact greater.
   D. MobiSmart's revenue by focusing on consumer AVs is not greater than that by focusing on long-haul trucks and it is in fact not greater.

6. When Toupet makes a hypothesis testing and concludes that _____, Type I error occurs
   A. MobiSmart's revenue by focusing on consumer AVs is greater than that by focusing on long-haul trucks but it is in fact not greater.
   B. MobiSmart's revenue by focusing on consumer AVs is greater than that by focusing on long-haul trucks and it is in fact greater.
   C. MobiSmart's revenue by focusing on consumer AVs is not greater than that by focusing on long-haul trucks but it is in fact greater.
   D. MobiSmart's revenue by focusing on consumer AVs is not greater than that by focusing on long-haul trucks and it is in fact not greater.

7. If Toupet concerned much more about the tremendous loss by focusing on long-haul trucks than that by focusing on consumer AVs, the alternative hypothesis would be _____. Given that u (LHT) represents the true market performance by focusing on long-haul trucks, and u (CAV) represents the true market performance by focusing on consumer AVs.
   A. Ha: u (LHT) > u (CAV)
   B. Ha: u (LHT) = u (CAV)
   C. Ha: u (LHT) < u (CAV)
   D. Ha: u (LHT) $\neq$ u (CAV)

接 次 頁

題號： 345
科目： 統計學(H)
節次： 2

國立臺灣大學 111 學年度碩士班招生考試試題

題號：345
共 9 頁之第 5 頁

8. If Toupet concerned much more about the tremendous loss by focusing on consumer AVs than that by focusing on long-haul trucks, the null hypothesis would be _____.?
   A. Ho: u (LHT) > u (CAV)
   B. Ho: u (LHT) = u (CAV)
   C. Ho: u (LHT) < u (CAV)
   D. Ho: u (LHT) $\neq$ u (CAV)

9. MobiSmart's marketing department found out that market performance (MP) and competitors' moves (CM) showed an exactly functional relationship, i.e., MP would be always identical to a constant minus CM multiplied by a positive coefficient without any exception. Which of the following is correct?
   A. MobiSmart's marketing department could run a regression of MP on CM to estimate the coefficient of CM.
   B. There is an error term following a normal distribution in the regression involving MP and CM.
   C. The covariance between MP and CM was negative.
   D. If the significance level was set extremely low, the lower limit of confidence interval of the estimated coefficient of CM would be reduced.

MobiSmart's marketing department collected five incumbent players' information on market performance (MP), partners' support (PS), and market potential in different segment (MPDS). The PS scores for five incumbent players were 357, 351, 352, 354, and 349, respectively, and MPDS were corresponding 89,985 units, 90,165 units, 90,135 units, 90,075 units, and 90,225 units, respectively. Two regressions were run. In the first regression, MP served as the dependent variable and PS served as the independent variable. In the second regression, the dependent variable was employed but the independent variable was replaced by MPDS. Please answer the following four questions.

10. Which of the followings regarding p-value is correct?
    A. The p-value in two regressions could not be compared unless the information of MP was provided.
    B. The p-value to the F-test in the first regression was different from that of the second regression, but the p-value to the t-test in the first regression was identical to that of the second regression.
    C. The p-values to the F-test and t-test in the first regression were different from those in the second regression.
    D. The p-values to the F-test and t-test in the first regression were identical to those in the second regression.

見背面

題號： 345
科目： 統計學(H)
節次： 2

國立臺灣大學 111 學年度碩士班招生考試試題

題號：345
共 9 頁之第 6 頁

11. Which of the followings regarding the regression coefficient is correct?
    A. The magnitude of regression coefficients in two regressions could not be compared unless the information of MP was provided.
    B. The coefficients in two regressions must be identical in magnitude but different in valence.
    C. The magnitude of regression coefficient in the first regression was smaller than that in the second regression.
    D. The magnitude of regression coefficient in the first regression was greater than that in the second regression.

12. Which of the followings regarding the standardized regression coefficient is correct?
    A. The magnitude of standardized regression coefficients in two regressions could not be compared unless the information of MP was provided.
    B. The standardized coefficients in two regressions must be identical in magnitude but different in valence.
    C. The magnitude of standardized regression coefficient in the first regression was smaller than that in the second regression.
    D. The magnitude of standardized regression coefficient in the first regression was greater than that in the second regression.

13. Which of the followings regarding R-square is correct?
    A. The information of MP should be provided in order to compare R-square between two regressions.
    B. R-square would be increased if significance level was relatively lower.
    C. R-square in two regressions were different.
    D. R-square in two regressions were identical.

14. An intern in MobiSmart's marketing department accidently merged PS and MPDS into a new independent variable X such that the first five observations of X were identical to PS and the remaining five observations of X were identical to MPDS. And the corresponding MP was not changed. That is, the first five observations of MP were identical to the remaining five observations of MP. Which of the followings is correct?
    A. The information of MP should be provided in order to estimate R-square and regression coefficients.
    B. R-square in the merged data would be larger than R-square in the first regression or in the second regression due to double size of observations.
    C. R-square and regression coefficient in the merged data would be 0 (or close to 0 due to calculation error).
    D. P-value to the F-test in the merged data would be 0 (or close to 0 due to calculation error).

接 次 頁

題號： 345
科目： 統計學(H)
節次： 2

國立臺灣大學 111 學年度碩士班招生考試試題

題號：345
共 9 頁之第 7 頁

The same intern would like to report the ANOVA table following the revenue forecasts made by MobiSmart's market department (Internal) and the contracted marketing research firm (External) for MobiSmart's strategic focus on consumer AVs (CAV), long-haul tracks (LHT), and both. The forecasted data was listed as follows.

Focus	Source	
	Internal	External
CAV	3.4	1.2
	2.8	.6
LHT	2.4	3.8
	2.0	3.4
Both	2.0	2.6
	1.8	2.2

Please fill in the following missing values.

ANOVA Output (Dependent Variable: Sales)

Source	df	SS	MS	F
Focus	"A"	"F"	"J"	"M"
Source	"B"	"G"	"K"	
Interaction	"C"	"H"	"L"	
Error	"D"	"I"	10.33	
Total	"E"	---		

15. "A" is equal to ____.

16. "B" is equal to ____.

17. "C" is equal to ____.

18. "D" is equal to ____.

19. "E" is equal to ____.

20. "F" is equal to ____.

21. "G" is equal to ____.

22. "I" is equal to ____.

見背面

題號： 345
科目： 統計學(H)
節次： 2

國立臺灣大學 111 學年度碩士班招生考試試題

題號：345
共 9 頁之第 8 頁

23. "J" is equal to ____.
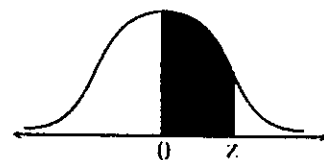
24. "K" is equal to ____.

25. "M" is equal to ____.

接 次 頁

題號： 345
科目： 統計學(H)
節次： 2

國立臺灣大學 111 學年度碩士班招生考試試題

題號：345
共 9 頁之第 8 頁

題號： 345
科目： 統計學(H)
節次： 2

國立臺灣大學 111 學年度碩士班招生考試試題

題號：345

共 9 頁之第 9 頁

常態分配表

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990
3.1	0.4990	0.4991	0.4991	0.4991	0.4992	0.4992	0.4992	0.4992	0.4993	0.4993
3.2	0.4993	0.4993	0.4994	0.4994	0.4994	0.4994	0.4994	0.4995	0.4995	0.4995
3.3	0.4995	0.4995	0.4995	0.4996	0.4996	0.4996	0.4996	0.4996	0.4996	0.4997
3.4	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4998
3.5	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998

試題隨卷繳回