

มหาวิทยาลัยสงขลานครินทร์

คณะวิศวกรรมศาสตร์

สอบปลายภาค : ภาคการศึกษาที่ 1

ปีการศึกษา : 2557

วันที่สอบ : 8 ธันวาคม 2558

เวลาสอบ : 13.30 - 16.30 น.

รหัสวิชา : 242-500

ห้องสอบ : xxx A401

ชื่อวิชา : ระเบียบวิธีวิทยาการวิจัยและพัฒนา (Research and Development Methodologies)

ผู้ออกข้อสอบ : พิชญ์ ตันฑิตย์

คำสั่ง: อ่านรายละเอียดของข้อสอบ และคำแนะนำให้เข้าใจก่อนเริ่มทำข้อสอบ

อนุญาต: เครื่องเขียนต่างๆ เช่น ปากกา หรือ ดินสอ, กระดาษโน้ตขนาด A4 เขียนด้วยลายมือตนเอง พร้อมชื่อ นามสกุลและรหัสนักศึกษา จำนวน 1 แผ่น (2 หน้า)

ห้าม: ห้ามนำข้อสอบทั้งหมดหรือบางส่วนออกนอกห้องสอบ และห้ามหยิบยืมหนังสือ เอกสาร เครื่องคำนวณ หรืออุปกรณ์อื่นๆซึ่งกันและกัน

เวลา: 3 ชั่วโมง (180 นาที)

คำแนะนำ:

- * ข้อสอบมี 15 หน้า 7 ข้อใหญ่ คะแนนรวม 160 คะแนน (25%) ให้ทำทุกข้อ และตอบเป็นภาษาไทย
- * อ่านคำสั่งในแต่ละข้อให้เข้าใจก่อนลงมือทำ
- * เขียนคำตอบลงในข้อสอบ คำตอบส่วนใดอ่านไม่ออก จะถือว่าคำตอบนั้นผิด
- * อย่าลืม เขียน ชื่อ-นามสกุล และรหัสนักศึกษา ลงในข้อสอบทุกแผ่น
- * แบนกระดาษโน้ตขนาด A4 ในข้อสอบ เมื่อส่งข้อสอบ

ทุกกรณีในการสอบ: มีโทษขั้นต่ำปรับตกในรายวิชานั้นและพักการเรียน 1 ภาคการศึกษา โทษสูงสุดให้ออก

ข้อที่	เวลา (นาที)	คะแนน	ได้	ข้อที่	เวลา (นาที)	คะแนน	ได้
1	35	30		5	25	25	
2	35	30		6	20	20	
3	35	30		7	10	10	
4	20	15		รวม		รวม	
รวม	180	160					25%

คำถามข้อที่ 1 หลักการพื้นฐานทางสถิติ

(30 คะแนน/35 นาที)

- a) จงยกตัวอย่างข้อมูลเชิงปริมาณ (quantitative data) หรือข้อมูลเชิงคุณภาพ (qualitative data) มาอย่างละ 2 ตัวอย่าง (4 คะแนน)

- b) จงจับคู่ตัวแปรแบบต่างๆเข้ากับระดับการวัดค่าตัวแปร (Scale of measurement) ต่อไปนี้ (3 คะแนน)

_____ interval scale

_____ nominal scale

_____ ordinal scale

A. ค่าดัชนีมวลกาย

น้อยกว่า 18.50

ระหว่าง 25 - 29.90

ระหว่าง 23 - 24.90

ระหว่าง 18.50 - 22.90

มากกว่า 30

B. ผลการวิเคราะห์เซลล์มะเร็งด้วยการประมวลผลภาพ

เป็นเซลล์มะเร็ง

ไม่ใช่เซลล์มะเร็ง

C. หน่วยความจำที่ใช้ในการประมวลผลเป็น MB

- c) จงเปรียบเทียบสถิติเชิงอนุมาน (Inferential Statistics) และสถิติเชิงพรรณนา (Descriptive Statistics) พร้อมยกตัวอย่างชื่อเทคนิคทางสถิติมาประเภทละ 3 ชื่อ (8 คะแนน)

d) จงอธิบายความหมายและการคำนวณหา degree of freedom (df) ทั้งในกรณีที่มีข้อมูลกลุ่มเดียวหรือมีข้อมูลมากกว่าหนึ่งกลุ่ม (2 คะแนน)

e) จงอธิบายแตกต่างระหว่างความคลาดเคลื่อนเชิงระบบ (Systematic Errors) และความคลาดเคลื่อนแบบสุ่ม (Random Errors) และยกตัวอย่างประกอบให้ชัดเจน พร้อมทั้งระบุแนวทางการแก้ไขความคลาดเคลื่อนดังกล่าว (6 คะแนน)

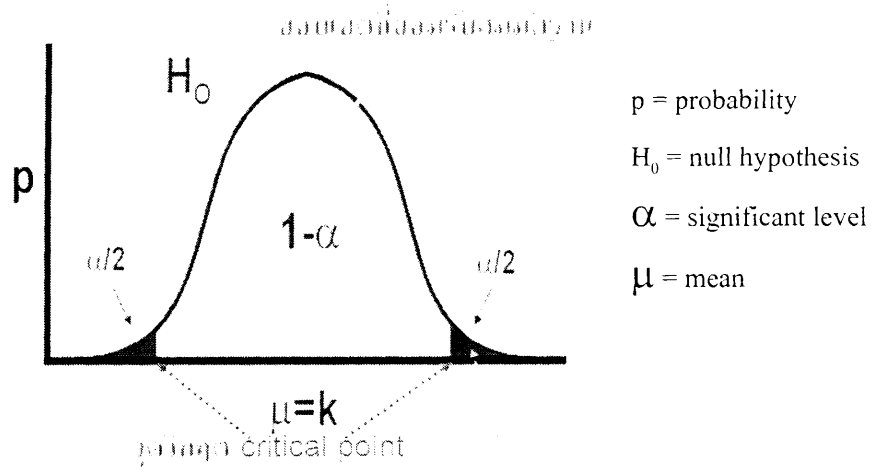
- f) จงยกตัวอย่างความผิดพลาดที่อาจเกิดขึ้นในการทดลอง การได้มาซึ่งข้อมูล และการวิเคราะห์สรุปทางสถิติอย่างน้อย 3 อย่าง (3 คะแนน)

- g) จงระบุประเด็นที่ควรพิจารณาในการตรวจสอบคุณภาพของเครื่องมืออย่างน้อย 4 อย่าง (4 คะแนน)

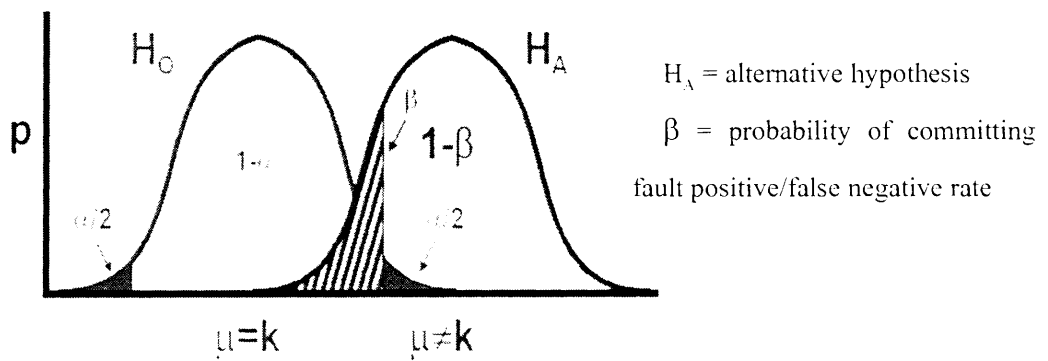
คำถามข้อที่ 2 ค่าทางสถิติในการตรวจสอบสมมติฐานหรือคุณภาพของเครื่องมืออุปกรณ์ต่างๆ (30 คะแนน/35 นาที)

- a) จงอธิบายความหมายและการนำไปใช้ในเชิงสถิติของการกระจายแบบปกติ (Normal Distribution) พร้อมวาดภาพประกอบ (5 คะแนน)

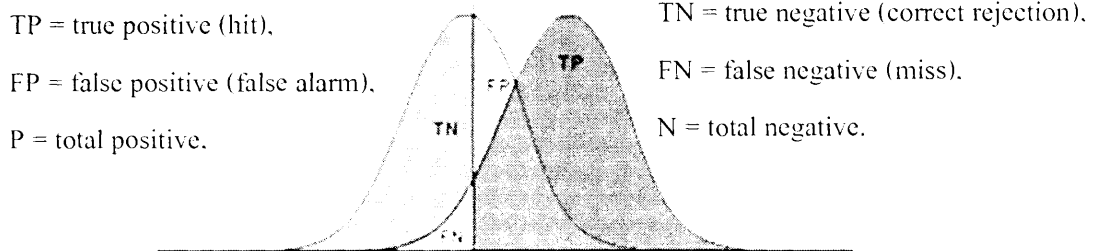
จงใช้ภาพและนิยามต่อไปนี้ ประกอบการอธิบายความหมายของค่าและสัญลักษณ์ทางสถิติต่าง ๆ ต่อไปนี้ ในข้อ b) ถึง i)



ภาพที่ 1



ภาพที่ 2



ภาพที่ 3

b) จงอธิบายว่าอำนาจในการทดสอบสมมติฐาน (Power) คืออะไร และจะเพิ่มอำนาจในการทดสอบสมมติฐานได้อย่างไร (5 คะแนน)

c) จงอธิบายความสัมพันธ์ระหว่าง False Negative และ False Positive (3 คะแนน)

d) จงยกตัวอย่างกรณีที่ควรให้ความสำคัญกับการปรับลด False Positive มากกว่า False Negative (2 คะแนน)

e) จงยกตัวอย่างกรณีที่ควรให้ความสำคัญกับการปรับลด False Negative มากกว่า False Positive (2 คะแนน)

f) จงอธิบายความแตกต่างระหว่างความเที่ยง (Precision) และความแม่นยำ (Accuracy) พร้อม
วาดภาพประกอบการอธิบาย (3 คะแนน)

g) จงอธิบายความหมายของความไว (Sensitivity) พร้อมยกตัวอย่าง (3 คะแนน)

h) จงอธิบายความหมายความจำเพาะ (specificity) พร้อมยกตัวอย่าง (3 คะแนน)

คำถามข้อที่ 3 สถิติแบบต่างๆ

(30 คะแนน/35 นาที)

จากข้อ a) ถึง c) จงระบุว่างานหรือคำถามเหล่านี้ ควรใช้วิธีการทางสถิติวิเคราะห์ข้อมูลเพื่อทดสอบ
สมมติฐานแบบใด

a) การทดลองเพื่อหาว่าความเร็วในการพิมพ์โดยใช้คีย์บอร์ดแบบธรรมดาและแบบ
Ergonomic แตกต่างกันหรือไม่ (2 คะแนน)

b) การทดลองเพื่อหาว่าอายุการใช้งานของฮาร์ดิสก์แต่ละยี่ห้อแตกต่างกันหรือไม่ (2 คะแนน)

c) การวิเคราะห์เพื่อทำนายปริมาณและแนวโน้มของการใช้งาน bandwidth ของเครือข่ายคอมพิวเตอร์ของภาควิชา ในแต่ละ ชั่วโมงของวัน แต่ละวันของสัปดาห์ แต่ละเดือนของปี (2 คะแนน)

d) ในการวิเคราะห์อนุกรมเวลา (Time Series Analysis) จงระบุและอธิบายองค์ประกอบของสูตรที่ใช้ในการพยากรณ์ (8 คะแนน)

- c) จงอธิบายความแตกต่างระหว่างการประมาณค่าในช่วง (Interpolation) และการประมาณค่า
นอกช่วง (Extrapolation) พร้อมทั้งวาดภาพประกอบ (4 คะแนน)

- f) จงยกตัวอย่างมาอย่างน้อย 3 กรณีที่ ควรปรับเปลี่ยนวิธีการหาค่าสหสัมพันธ์ (correlation)
ใหม่ พร้อมอภิปรายว่าเพราะเหตุใดและควรทำอย่างไร (6 คะแนน)

- g) จงระบุการนำไปใช้ประโยชน์ของวิธีการทางสถิติในการวิเคราะห์การถดถอย (Regression
Analysis) (2 คะแนน)

- h) จงอธิบายที่มาและความหมายของ Over-fitting และแนวทางการแก้ปัญหาโดยใช้ validation set (4 คะแนน)

คำถามข้อที่ 4 การสุ่มตัวอย่างทางสถิติ

(15 คะแนน/20 นาที)

จงตอบคำถามต่อไปนี้

- i) จงระบุข้อดีและข้อเสียของการกำหนดจำนวนตัวอย่าง (sample) ในการสุ่ม (sampling) น้อยหรือมากเกินไป (2 คะแนน)

- j) การสุ่มตัวอย่างหรือการคัดเลือกแบบไม่รู้โอกาสในการเลือก (*Probability Sampling*) และการคัดเลือกตัวอย่างแบบรู้โอกาสในการเลือก (*Non-probability Sampling*) แตกต่างกันอย่างไรร (5 คะแนน)

จงระบุข้อควรระวังในการสุ่มตัวอย่างแบบต่างๆต่อไปนี้ พร้อมยกตัวอย่างประกอบ

k) การสุ่มอย่างง่าย (Simple Random Sampling) (2 คะแนน)

l) การสุ่มอย่างเป็นระบบ (Systematic Sampling) (2 คะแนน)

จากการสุ่มตัวอย่างแบบต่างๆต่อไปนี้ จงยกกรณีตัวอย่างให้พอเห็นภาพ

m) การสุ่มแบบแบ่งชั้น (Stratified Sampling) (2 คะแนน)

n) การสุ่มแบบแบ่งกลุ่ม (Probability Proportional to Size Sampling) (2 คะแนน)

คำถามที่ 5 การออกแบบการทดลองและโมเดลคอมพิวเตอร์ (25 คะแนน/25 นาที)

a) จงอธิบายความสำคัญของ Randomization ที่มีต่อการออกแบบการทดลอง (4 คะแนน)

b) จงอธิบายความสำคัญของ Replication ที่มีต่อการออกแบบการทดลอง (4 คะแนน)

c) จงอธิบายความสำคัญของ Blocking ที่มีต่อการออกแบบการทดลอง (4 คะแนน)

d) จงอธิบายความแตกต่างระหว่าง replicates และ repeated measurements (4 คะแนน)

e) ความแม่นยำของโมเดลจำลองคอมพิวเตอร์ขึ้นอยู่กับอะไรบ้าง (4 คะแนน)

f) จงอธิบายความหมายของ absolute zero และการเปรียบเทียบความแตกต่างระหว่างสองระดับใน interval data และ ratio data (5 คะแนน)

คำถามที่ 6 Chaos and Reliability (20 คะแนน/20 นาที)

a) ระบบ Chaos มีพฤติกรรมอย่างไร (4 คะแนน)

b) เหตุใดเราจึงไม่สามารถใช้เอาต์พุตของระบบป้อนกลับไปเป็นอินพุตของระบบเพื่อหาค่าเอาต์พุตตัวใหม่ได้ในระยะยาว (4 คะแนน)

c) จงอธิบายความแตกต่างระหว่างความน่าเชื่อถือ (Reliability) ของระบบแบบอนุกรมและแบบขนาน (4 คะแนน)

d) จงอธิบายความแตกต่างในการเพิ่มความน่าเชื่อถือ (Reliability) ของฮาร์ดแวร์และซอฟต์แวร์ (8 คะแนน)

คำถามที่ 7 General Topics

(10 คะแนน/10 นาที)

จงระบุว่าประโยคต่อไปนี้จริง (T) หรือเท็จ (F)

- a) _____ วิธีการสุ่มแบบแบ่งกลุ่ม (Probability Proportional to Size Sampling) ทำให้บางประชากรไม่มีโอกาสได้รับเลือกเลย
- b) _____ วิธีการสุ่มแบบหลายชั้นตอน (Cluster and Multistage Sampling) ทำให้ประชากรมีโอกาสดังกล่าวได้รับเลือกเท่ากัน
- c) _____ การเลือกเอาข้อมูลแบบตามสะดวกหรือโดยบังเอิญ (Accidental/Convenient Sampling) อาจมีปัญหากรณีข้อมูลที่ได้อาจไม่เกี่ยวข้องหรือไม่สามารถใช้ในการตอบคำถามวิจัยเลยก็ได้
- d) _____ การเลือกรูปแบบการวิจัยขึ้นอยู่กับวัตถุประสงค์ของการวิจัย ข้อมูลหรือตัวแปรในการวิจัย และความเกี่ยวข้องซึ่งกันและกันของตัวแปร
- e) _____ การเปลี่ยนข้อมูล (data) เปลี่ยนไปเป็นข่าวสาร (information) เริ่มตั้งแต่การเก็บรวบรวมข้อมูล นำมาเรียบเรียง นำเสนอ จัดการและวิเคราะห์ จากนั้นก็สรุปและแปลความหมาย ให้ได้สาระของข่าวสารที่ต้องการ
- f) _____ หากไม่สามารถรวมปัจจัยที่เกี่ยวข้องได้ทั้งหมด การทดลอง/ศึกษานั้นๆอาจจะมีอคติหรือความลำเอียง (bias)
- g) _____ ในการออกแบบและทำการทดลอง นิยม/ควรให้มีผู้ดำเนินการเป็นคนเดียวกัน เพื่อป้องกันความสับสน
- h) _____ การจับสลากหรือการทอดลูกเต๋าเป็นการสุ่มที่ประชากรมีโอกาสดังกล่าวได้รับเลือกไม่เท่ากัน
- i) _____ การวางแผนก่อนการทดลองเป็นปัจจัยที่สำคัญในการที่จะทำการทดลองให้ได้ผลสำเร็จ
- j) _____ การทดสอบสมมติฐานในทางสถิติคือวิธีการตัดสินใจโดยใช้ข้อมูลทางสถิติที่สรุปได้จากการทดลอง เพื่อตัดสินใจว่าผลที่ได้มีข้อมูลข่าวสารเพียงพอที่จะก่อให้เกิดความสงสัยที่มีต่อทฤษฎีหรือความรู้ที่เข้าใจกันมาดั้งเดิม

----End of Examination----

Pichaya Tandayya Lecturer