**ใบงานที่ 7**

**บรรทัดที่ 1**

**บรรทัดที่ 2**

**บรรทัดที่ 3**

**1.บรรทัดที่ 1 : ความหมายของข้อมูลตัวแปรแต่ละประเภท พร้อมตัวอย่าง 5 ลักษณะ**

**ตอบ**  ความหมายและลักษณะของตัวแปรตามทั้ง 5 ประเภท มีดังนี้

**1.ตัวแปรแบบต่อเนื่อง (Continuous outcome)** หมายถึง ตัวแปรที่สามารถแปรค่าได้ต่อเนื่องกันภายในช่วงใดช่วงหนึ่งหรือสามารถมีค่าเป็นเลขทศนิยมได้ และสามารถแสดงลําดับที่ได้คือ ค่าที่มากกว่าแสดงว่ามีคุณสมบัติของตัวแปรมากกว่าค่าที่น้อยกว่า

**ตัวอย่างเช่น** ความสูง นํ้าหนัก ราคาสินค้า ความถนัด สติปัญญา ผลการเรียน ซึ่งตัวแปรเหล่านี้สามารถมีค่าเป็นเลขทศนิยมได้ เช่น ความสูง 175.8 ซม. นํ้าหนัก 15.5 ก.ก.

**2.ตัวแปรแบบแจงนับ (Categorical outcome)** หมายถึง ตัวแปรที่บอกคุณลักษณะที่แตกต่างกันเป็นประเภทๆ ตัวแปรนี้อาจเรียกตามจำนวนประเภทที่แบ่งไว้ได้แก่ “ตัวแปรสองประเภท” หรือ “ตัวแปรคู่” (dichotomous variable)  **ตัวอย่างเช่น** เพศ แบ่งเป็นหญิงและชาย และ “ตัวแปรหลายประเภท” (polytomous variable) เช่น สถานภาพสมรส แบ่งเป็น โสด สมรส ม่าย หย่า แยก

**3. Count** หมายถึง การวัดอัตราของตัวแปรตามในกลุ่มควบคุมในช่วงเวลาหนึ่ง

**ตัวอย่างเช่น** อัตราการหายของสิวของกลุ่มที่ทากับไม่ทาขมิ้นชัน, ความเสี่ยงในการป่วยของคนที่มีปัจจัยเสี่ยงกับคนที่ไม่มีปัจจัยเสี่ยง , ความแตกต่างของการอัตราตายของคนในเขตเมืองกับเขตชนบท,

**4. Survival** หมายถึง เทคนิควิธีวิเคราะห์ทางสถิติที่ใช้ศึกษาเกี่ยวกับเวลาตั้งแต่เริ่มต้นศึกษาเหตุการณ์หรือสิ่งที่สนใจไปจนกระทั่งเกิดเหตุการณ์ที่ต้องการศึกษา เพื่อให้ทราบว่าเกิดเหตุการณ์นั้นหรือไม่ (whether) เมื่อไร (when)

**ตัวอย่างเช่น** การศึกษาภาวะแทรกซ้อนในผู้ป่วยเบาหวาน, การศึกษาอัตราการตายของผู้ป่วยเอดส์

**2.บรรทัดที่ 2 : หมายถึงอะไร พร้อมยกตัวอย่างคำถามวิจัยและสถิติที่ใช้วิเคราะห์เพื่อให้ได้ค่าตามบรรทัดที่ 2**

**ตอบ** ค่าที่อยู่ในบรรทัดที่ 2 หมายถึงค่าทางสถิติ ซึ่งเป็นค่าที่คำนวณได้จากข้อมูลที่เป็นตัวอย่าง เป็นตัวที่บ่งชี้ ถึงคุณลักษณะของตัวอย่าง ซึ่งแต่ละค่ามีความหมายดังต่อไปนี้

 **1. ค่าเฉลี่ยเลขคณิต (Arithmetic Mean)** ค่าเฉลี่ยเลขคณิตเป็นค่าเฉลี่ยที่นิยมใช้กันมากที่สุด จะแทนด้วย μ (มิว) เมื่อคำนวณจากข้อมูลทั้งประชากร และแทนด้วย $\overbar{x}$ เมื่อคำนวณจากข้อมูลตัวอย่าง การ

คำนวณหาได้จากผลรวมของข้อมูลทุกค่าแล้วหารด้วยจำนวนข้อมูลทั้งหมด

 **ตัวอย่างคำถามวิจัย:** จงหาค่าเฉลี่ยรายได้ของหัวหน้าครอบครัวใน Stat village

 **สถิติที่ใช้ในการวิเคราะห์หาค่า Mean :** t test หรือ Z test

 **2. ค่ามัธยฐาน (Median)** เป็นค่าที่บอกภาพรวมของข้อมูล โดยพิจารณาจากตำแหน่งกลางของข้อมูลที่เรียกลำดับจากน้อยไปหามากแทนด้วย Me

 **3.สัดส่วน (Proportion)** คือ ความสัมพันธ์ของจำนวนย่อยกับจำนวนรวมทั้งหมด กล่าวคือ ให้ถือจำนวนรวมทั้งหมดเป็น 1 ส่วน

**ตัวอย่างคำถามวิจัย:** จงหาอัตราความชุกของการเจ็บป่วยของประชากรใน Stat village

**สถิติที่ใช้ในการวิเคราะห์หาค่า Proportion :** Z test หรือ Binomial exact probability test

**4.ค่ามัธยฐานระยะปลอดเหตุการณ์ (Median survival time)** คือ ระยะเวลาจนกระทั่งเกิดเหตุการณ์ เช่น แทนที่จะดูว่าผลการรักษาว่ามีตายหรือไม่ ก็มาดูว่ายืดอายุไปได้ (รอดชีพ) นานเท่าใด

**ตัวอย่างคำถามวิจัย :** จากการศึกษาประสิทธิผลของขมิ้นชันในการรักษาสิวโดย Randomized controlled trial กลุ่มละ 10 คน กลุ่มหนึ่งให้ทาขมิ้นชันทุกวันในสามวัน ส่วนอีกกลุ่มหนึ่งล้างหน้าตามปกติโดยไม่ทาขมิ้นชัน ระยะเวลาของการศึกษาเท่ากับ 14 วัน การหายหมายถึงสิวยุบจนมองไม่เห็นร่องรอยภายใน 2 สัปดาห์ อยากทราบว่าอัตราการหายของสิวภายใน 2 สัปดาห์มีค่าเท่าใด

**สถิติที่ใช้ในการวิเคราะห์หาค่า Median survival time :** วิธี Kaplan-Meier

**3.บรรทัดที่ 3 : หมายถึงอะไร พร้อมยกตัวอย่างคำถามวิจัยและสถิติที่ใช้วิเคราะห์เพื่อให้ได้ค่าตามบรรทัดที่ 3**

**ตอบ**  บรรทัดที่ 3 หมายถึง วิธีการวิเคราะห์ข้อมูลทางสถิติซึ่งมีดังต่อไปนี้

 **1.linear regression หรือ การวิเคราะห์การถดถอย** คือ การศึกษาถึงข้อมูลที่ประกอบด้วยตัวแปร2 ตัว หรือ เรียกว่า ตัวแปรคู่ (bivariate data)โดยการวิเคราะห์จะนำตัวแปรทั้งสองมาพิจารณาพร้อม ๆ กัน เพื่อศึกษาถึงความสัมพันธ์ระหว่างตัวแปร 2 ตัว หรือ 2 ลักษณะโดยที่ต้องทราบค่าของตัวแปรตัวหนึ่ง หรือต้องกำหนดค่าของตัวแปรตัวหนึ่งไว้ล่วงหน้า โดยการวิเคราะห์นี้จะให้ค่าสัมประสิทธิ์การถดถอย

 **1.1 simple linear regression** เป็นการวิเคราะห์ความสัมพันธ์ในเชิงคาดคะเนที่ตัวแปรตามเป็นตัวแปรต่อเนื่องตัวแปรเดียวโดยที่ตัวแปรต้นเป็นตัวแปรประเภทใดก็ได้ แต่ตัวแปรต้นต้องมีเพียงหนึ่งตัว

 **ตัวอย่างคำถามวิจัย:**จากการศึกษาข้อมูลของนักศึกษาในมหาวิทยาลัยแห่งหนึ่งจำนวน 143 คน จงศึกษาว่าน้ำหนักสามารถใช้ทำนายอัตราชีพจรขณะพักได้หรือไม่อย่างไร

**1.2 multiple linear regression** การศึกษาความสัมพันธ์ระหว่าง ตัวแปรตาม 1 ตัว กับตัวแปรอิสระมากกว่าหนึ่งตัว

**ตัวอย่างคำถามวิจัย:**ในการจะศึกษาประสิทธิภาพการใช้น้ำมันของรถยนต์ เราไม่สามารถจะเอาขนาดของเครื่องยนต์มาเป็นตัวกำหนดเพียงอย่างเดียว จะต้องคำนึงถึงน้ำหนักตัวรถ น้ำหนักคนขับ อายุของเครื่องยนต์ ความเสียดทานต่อผิวถนนของล้อรถ พูดง่ายๆคือหากต้องการพยากรณ์อัตราความสิ้นเปลืองของน้ำมันเชื้อเพลิง หรืออัตราการใช้น้ำมัน (กิโลเมตร/ลิตร) แล้วจะต้องคำนึงถึงตัวแปรอิสระมากกว่าหนึ่งตัวแปร

 **2.Logistic regression** เป็นการทำนายว่า จะเกิดเหตุการณ์หนึ่งขึ้นหรือไม่หรือมี โอกาสเกิดขึ้นมากน้อยเพียงใด โดยมีการกำหนดค่าตัวแปรตัวหนึ่งหรือหลายตัวที่คาดว่าจะส่งผลต่อการเกิดเหตุการณ์นั้นๆ และในที่สุดก็จะทำให้เราเข้าใจสาเหตุการเกิดเหตุ การณ์นั้นๆได้ในที่สุด

            **ตัวอย่างคำถามวิจัย :** ปัจจัยด้าน อายุ ดัชนีมวลกาย จำนวนการสูบบุหรี่ ประวัติการป่วยด้วยโรคเบาหวาน ความดันโลหิต  เป็นปัจจัยเสี่ยงต่อการเกิดโรคหัวใจหรือไม่

**3. Possion regression** หมายถึง เป็นการศึกษาเกี่ยวกับความสัมพันธ์ระหว่างตัวแปร ซึ่งมีความสัมพันธ์เป็นตัวแปรแบบการถดถอยปัวซง (Poisson regression modal) เพื่อใช้ในการทำนายหรือพยากรณ์ค่าของตัวแปรที่ศึกษา

**ตัวอย่างคำถามวิจัย :** จำนวนครั้งของการเจ็บป่วยของหัวหน้าครอบครัวในหมู่บ้าน stat village ในปีที่ผ่านมา

**4. Cox regression** คือ การวิเคราะห์การอยู่รอดขั้นสูง ใช้ในการศึกษาความสัมพันธ์ระหว่างตัวแปรตามที่สนใจ กับตัวแปรอิสระต่างๆ ตั้งแต่หนึ่งตัวขึ้นไปโดยเป็นข้อมูลที่มี ช่วงเวลาเข้ามาเกี่ยวข้องและมีลักษณะเป็นข้อมูลไม่สมบูรณ์

 **ตัวอย่างคำถามวิจัย :** อัตราการอยู่รอดของผู้ป่วยโรคมะเร็งทุกชนิดในปี พ.ศ. 2535-2555