|  |  |  |
| --- | --- | --- |
| |  | | --- | | **Description: http://virtualcampus.pupr.edu/images/ci/sets/set01/lesson_on.gifModulo 06: Medidas de Tendencia Central y Variabilidad** |  |  | | --- | | **ENGI 2210/2270 Probability and Statistics for Engineers** | |
|  |
| **Modulo 06: Medidas de Tendencia Central y Variabilidad** |
| **🗒Introducción** |
| **Durante este modulo se presentaran las medidas de tendencia central (valor alrededor del cual están la mayoría de los datos de un experimento) . Se presentara el cómputo de los cuartiles. También las medidas de variabilidad (cuan diferentes son los datos de un experimento).**  **Medidas de singular importancia en el análisis de los datos y de uso continuo y cotidiano en el mundo comercial y en la toma de decisiones.** |
| **☑ Learning Goals /Outcomes (What you need to know)** |
| **Al completar este modulo el estudiante será capaz de :**   * **Calcular Medidas de tendencia Central:**   + **Promedio**   + **Moda**   + **Mediana.** * **Conocer definición de cuartiles y su computo** * **Calcular Medidas de Dispersión:**   + **Rango**   + **IQR**   + **Varianza**   + **Desviación Estándar**   + **Coeficiente de variación**   **Leccion#6.1: Medidas de Tendencia Central**  **Leccion#6.2: Cuartiles y Percentiles**  **Leccion#6:3 Medidas de variabilidad o Dispersión** |
| **🗍 Lección# 6.1 Medidas de Tendencia Central** |
| **Medidas de tendencia central- son medidas que nos indican alrededor o en la vecindad de que valor numérico, se hallan la mayoría de los datos de un experimento o estudio estadístico**  **En esta lección se estudiaran las siguientes medidas de tendencia central:**   * **Promedio (o la Media Aritmética)** * **Mediana** * **Moda** * **Rango Medio** * **Eje-Medio**   **Recordemos que una medida de tendencia central es aquel valor que representa el centro de los datos, ósea aquel valor donde la mayoría de los datos están a la derecha e izquierda del mismo. Son las medidas como Promedio, Mediana o Q2, Eje Medio y Rango Medio los que nos indican donde está el centro de los datos de un experimento o estudio.**  **Promedio: o Media Aritmética = La suma de todos los datos de un experimento divididos por la cantidad de datos sumados (tamaño de muestra o población según sea el caso)**   * **Poblacion** * **Muestra**   **Mediana (Md) = Q2 = P50 = (La mediana y Cuartil 2 y la Percentila 50 son lo mismo) = aquel valor de los datos ordenados donde el 50% de los datos son menores o iguales a el. Favor de estudiar más adelante el cómputo de Q2.**  **Moda (Mo) = Representa el numero que más se repite. Es de poca utilidad en muestras pequeñas. Es de alguna utilidad en la descripción de poblaciones o muestras bien grandes.**  **Por lo general más de cuatro modas en una muestra decimos que los datos no tienen moda.**  **OTRAS MEDIDAS DE TENDENCIA CENTRAL**  **Rango Medio = (Xmin+ Xmax)/2 = la suma del valor mínimo mas el máximo dividido por dos.**  **Eje-Medio= (Q3+Q1)/2 = la suma del cuartil 3 mas cuartil 1 dividido por dos.**  **EJEMPLO:**  **NOTA: En este ejemplo se presentara el computo de cuartiles utilizando la mediana como criterio. (Un poco de sentido común, en la próxima lección se presentara el computo utilizando formulas como lo presente el libro de texto)**  **Para los siguientes datos calcule: Promedio, Moda, Mediana, Eje Medio y Rango Medio.**  **DATOS: 2, 4, 4, 1, 1, 5, 3, 2, 1, 4, 5, 5**  **Hay 12 datos (n=12). Por lo tanto 12 posiciones \_1ra \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_12va\_**  **Como es necesario computar los cuartiles se deben ORDENAR los datos.**  **DATOS ORDENADOS: 1, 1, 1, 2, 2, 3, 4, 4, 4, 5, 5, 5**    **● Promedio= Σx/n = (1+1+1+2+2+3+4+4+4+5+5+5)/12 = 37/12 = 3.0833**   * **Moda = 4 y 5.** * **Computo de los cuartiles: Para determinar cuartiles se comienza con Q2 o la mediana. Una vez ordenados los datos buscamos en qué posición podemos dividir los datos en dos partes iguales, para 12 datos es lógico que sea entre medio de la posición 6ta y 7ma (6 posiciones a la izquierda y 6 posiciones a la derecha) Por lo tanto se calcula un promedio entre los valores que ocupan las anteriores posiciones ósea (3+4)/2=3.5 será la mediana Md=Q2=3.5**   **1, 1, 1, 2, 2, 3, 4, 4, 4, 5, 5, 5**  **|**  **Q2**   * **Determinada ya la mediana (Md) se procede a calcular Q1. Repitiendo el proceso anterior pero con los valores que quedan a la izquierda de la mediana. Ósea para 1, 1, 1, 2, 2, 3 ¿Cuál es su mediana? Otra vez por lógica observacional determinaremos que el punto medio está entre la 3ra y 4ta posición de los datos ordenados por lo que nuevamente sacamos un promedio Q1=(1+2)/2=1.5**   **1, 1, 1, 2, 2, 3, 4, 4, 4, 5, 5, 5**  **| |**  **Q1 =1.5 Q3=4.5**   * **Por último se vuelve a calcular la mediana de los restantes datos 4, 4, 4, 5, 5, 5 , la misma se hallara entre la 9na y 10ma posición. Ósea un promedio entre Q3=(4+5)/2=4.5** * **Resultado: Q2=(3+4)/2=3.5 Q1=(1+2)/2=1.5 Q3=(4+5)/2=4.5** * **Eje Medio=(Q1+Q3)/2= (1.5+4.5)/2=6/2=3** * **Rango Medio=(Xmin +Xmax)/2=(1+5)/2=3**   **En resumen para los datos : 1, 1, 1, 2, 2, 3, 4, 4, 4, 5, 5, 5, Md=3.5, =3.0833, Rango Medio=3.0, Eje Medio=3.0 estas medidas indican que la mayoría de los datos están en la vecindad del numero 3.0** |
|  |
| **🗍 Lección# 6.2 Cuartiles y Percentiles**  **Cuartil –dividir los datos ordenados en 4 partes iguales.**  **Percentil- Dividir los datos ordenados en 100 partes iguales**  **Cuartil1 =Q1= aquel número de los datos ORDENADOS donde el 25% de los datos son menores o iguales a el. Q1=P25 el cuartil uno es igual a la percentil 25.**  **Cuartil2=Q2=Md=P50 = aquel número de los datos ORDENADOS donde el 50% de los datos son menores o iguales a el. Q2=Md=P25 Cuartil dos es igual a la mediana que es igual al percentil 50.**  **Cuartil3=P75= aquel número de los datos ORDENADOS donde el 75% de los datos son menores o iguales a el. Q3=P75 Cuartil tres es igual a percentil 75.**  **FORMAS DE COMPUTAR:**   * **Método de Análisis Visual**   + **Se ORDENAN los datos**   + **Se calcula el punto medio de los datos ordenados ósea se calcula Q2. Si la muestra es impar Q2 será un dato observado. Si la muestra es par Q2 se calcula mediante el computo de un promedio de los datos observados al dividir los datos ordenados que están a la izquierda y derecha del punto que el 50% están a la izquierda y derecha.**   + **Se repite el procedimiento anterior calculando la mediana de los datos que están a la izquierda Q1. Luego se hace lo mismo calculando la mediana de los datos a la derecha Q3** * **Mediante Formulas de Posición en Datos Ordenados**   + **Posición de Q1=1/4 \*(n+1) , no da numero entero se redondea a entero menor si decimal es 0,1,2,3,4**   **Redondear a entero mayor si decimal es 6,7,8,9**  **Si el decimal es es .5 no podemos redondear posición .Se debe localizar esta posición dentro de los números ordenados.**  **Computo del Cuartil:**  **Se tiene que sacar un promedio entre el número que este en la posición a la izquierda y derecha donde se localizo valor con decimal .5 este número resultado del promedio será el valor del cuartil**   * **Posición de Q2= 2/4(n+1) Aplica redondeo anterior con una excepción si el decimal es .5 no se puede redondear aplica procedimiento antes presentado.** * **Posicion de Q3=3/4(n+1) Aplica redondeo anterior con una excepción si el decimal es .5 no se puede redondear aplica procedimiento antes presentado.**   **RECUERDE: LA FORMULA LE DA LA POSICION DENTRO DE LOS DATOS ORDENADOS PARA ENTONCES BUSCAR EL VALOR DEL DATO QUE SE ENCUENTRA EN LA POSICION. ES ESE VALOR EL CUARTIL . Recuerde posición NO ES el cuartil.**   * **Mediante Programa de Computadora. Favor de leer comentarios al finalizar el modulo sobre esta situación.**   **FORMULA PARA COMPUTO DE PERCENTIL**  **La percentil *p* es aquel valor de los datos ordenados donde el *p* por ciento de los datos son memores o iguales a el.**  **Para el computo del percentil *p* hacer los siguientes pasos:**  **1. Ordenar los datos de menor a mayor.**  **2. Calcular el índice *i* este nos indica la posición relativa en donde se hallara el valor de la data que representa el percentil deseado.**    ***p* – percentil deseado Por ejemplo percentil 85 , p=85**  **n – Tamaño de muestra o cantidad de datos.**  **i – posición de los datos ordenados donde se podrá localizar el valor**  **Percentil deseado.**  **3. Si *i* no es un entero , redondee hacia arriba al próximo entero. Si *i* es un entero , la**  **la percentil estará entre la posición *i e* i+1  *de los datos ordenados por lo que se***  **tendrá que sacar un promedio entre el valor que ocupe posición *i* y el valor que este**  **en la posición *i+1*.**  **EJEMPLO: Halle la percentil 85 para los siguientes datos:**  **23 34 34 35 35 35 35 37 37 38 38 39 39 40 40 40**  **Tenemos n=16 datos y están ordenados . (En esta ocasión no se necesita ORDENAR los datos)**  **Si se desea conocer quién es el percentil 85 , i= 85/100 \* 16 =13.6 el valor de la posición no es un numero entero se procede a redondear sl entero mayor i=14 por lo tanto el valor que se encuentre en la posición 14 es la percentila 85, en este ejemplo es el numero 40.**  **23 34 34 35 35 35 35 37 37 38 38 39 39 40 40 40 Dato experimental**  **1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 posición de dato ordenado**  **|**  **Posicion 14 Percentil 85 es el 40**  **Si se desea conocer quien es percentil 50, i=50/100\*16=8 como dio entero la posición es un promedio entre la posición i y i+1 osea (8+9)/2 = 8.5 esto significa que el percentil esta entre dos valores por lo que hay que sacar un promedio entre estos valores (37+37)/2 =37 es el valor del percentil 50.**  **23 34 34 35 35 35 35 37 37 38 38 39 39 40 40 40 Dato experimental**  **1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 posición de dato ordenado**  **|**  **Posicion 8.5 Percentil 85 es el promedio (37+37)/2=37**  **Es mucho mas fácil y preciso el utilizar el método visual antes expuesto para computar P25=Q1, P50=Q2=Md y P75=Q3.** |
| **🗍 Lección # 6.3 Medidas de Dispersión o Variabilidad**  **Dispersión o Variabilidad es cuan diferentes son los datos entre si. Si todos los datos son de igual magnitud ósea son iguales no hay variabilidad o esta será cero.**  **Variabilidad cuan diferentes son los datos. Dispersión cuan separados están los datos. Las siguientes medidas nos cuantifican el nivel de diferencia de los datos.**  **En el caso de la varianza, desviación estándar y coeficiente de variación estas medidas nos ofrecen cuan diferentes son los datos respecto a su promedio.**   * **Calcular Medidas de Dispersión:**   + **Rango**   + **IQR**   + **Varianza**   + **Desviación Estándar**   + **Coeficiente de variación**   **Rango= Xmax – Xmin = es la diferencia entre el valor máximo y el minimo**  **Rango Intercuartilico (Interquartile range)**   * **IQR = Q3 –Q1 es la diferencia entre el cuartil3 menos el cuartil 1.**   **Varianza**  **La varianza es una medida de cuan diferente son los datos de su promedio. Esta medida nos tiende a promediar la suma (para todos los datos) del cuadrado de la diferencia entre el dato y su promedio.**  **Al igual que el promedio la varianza se computa para:**   * **POBLACION** * **MUESTRA**     **El siguiente ejemplo nos ilustra el computo de la varianza. Calcule la varianza para la siguiente muestra:**  **DATOS: 1, 1, 2, 2, 3, 3**   * **Primeramente hay que calcular el promedio= ΣX/n=12/6=2** * **Luego se hace la siguiente tabla**  |  |  |  | | --- | --- | --- | | **X** | **(X - )** | **(X - )2** | | **1** | **1-2=-1** | **1** | | **1** | **1-2=-1** | **1** | | **2** | **2-2=0** | **0** | | **2** | **2-2=0** | **0** | | **3** | **3-2=1** | **1** | | **3** | **3-2=1** | **1** |   **ΣX = 12 Σ(X - ) = 0 Σ(X - )2 = 4**   * **Utilizando los resultados de la ultima columna de la tabla, se dividen por n-1**   **S2 = Σ(X - )2 /n-1 = 4/6-1 = 4/5 = 0.8**   * **La varianza FORMULA COMPUTACIONAL= es una simplificación matemática mediante la manipulación de la ecuación original para facilitar el computo manual de la varianza.**     **EN FORMULA COMPUTACIONAL**   * **RECORDAR QUE  NO SON IGUALES** * **EJEMPLO: Para X=1,2,3**   **14 ≠ 36**   * **Siempre las varianzas y desviaciones estándares son POSITIVAS. Nunca pueden ser negativas.**   **Desviación Estándar - por definición es la raíz cuadrada de la varianza. Dado que la varianza es el promedio la suma de los cuadrados de la diferencia entre el dato y su promedio, La desviación estándar es la raíz cuadrada de la varianza. Representando un valor más correcto del promedio de las diferencias.**  **Población**  **Muestra**  **Ejemplo: Utilizando el computo anterior de la varianza S2=0.8 por lo tanto**    **Coeficiente de variación se utiliza para conocer entre dos o más experimentos cuál de**  **ellos tiene mayor variabilidad.**  **Población**  **Muestra**  **EJEMPLO:**  **Determine cuál de los dos conjuntos de datos posee mayor variabilidad.**   |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **Data1** | **12** | **12** | **20** | **21** | **22** | **24** | **25** | **25** | **30** | **30** | **40** | | **Data2** | **8** | **8** | **10** | **10** | **15** | **15** | **16** | **20** | **20** | **23** | **25** |   **Se calcula el coeficiente de variación para ambos conjuntos de datos.**  **DATA1**  **Sample Size, n: 11**  **Mean: 15.45455**  **Variance, s^2: 36.07273**  **St Dev, s: 6.006058**  **Coeff. Of Var. CV= 6.006058/15.45455 x 100=38.86%**  **DATA2**  **Sample Size, n: 11**  **Mean: 23.72727**  **Variance, s^2: 64.61818**  **St Dev, s: 8.038544**  **Coeff. Of Var. CV=8.038544/23.72727 x 100=33.87%**  **El conjunto de datos DATA1 tiene mayor variabilidad por tener 38.86% vs 33.87% de DATA2** |
| **Conclusion (Closing statement paragraph)** |
| **En este modulo se presentaron las medidas de tendencia central (valor alrededor del cual están la mayoría de los datos de un experimento) , el computo de cuartiles y las medidas de variabilidad (cuan diferentes son los datos de un experimento).**  **Medidas que son necesarias para el análisis y comparación de los datos de un experimento. Estas medidas nos ayudan a la evaluación y decisión entre alternativas, para la toma de decisiones. Estas medidas se utilizan ampliamente durante el resto del curso.** |
| **🗁 Learning Activities (What you need to do)** |
| **Activities for this lesson:**  **Universidad Politécnica de Puerto Rico**  **EJERCISIOS DE PRACTICA**  **ENGI 2210/2270 Probability and Statistics for Engineers**  **Nombre:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Num:Est:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**  **Prof. Ing. José Raúl Díaz, PE;**    **I. Realicé el siguiente problema:**  **a- Para los siguientes datos: (Ordenados)**  **3 4 5 8 8 10 12 12 20 22**  **a) Calcule**  **1) la media o promedio 6) el rango**  **2) la mediana 7) el rango intercuartilico**  **3) la moda 8) la varianza**  **4) el rango medio 9) la desviación estándar**  **5) el eje medio 10) el coeficiente de variación** |
| **📋 Assignment #1** |
| **(Include due dates, time frames and point possible).**  **1. Para los siguientes datos:**  **1, 1, 2, 3, 3, 3, 3, 3, 4, 4, 5, 5, 5**  **a) Calcule**  **1) la media o promedio**  **2) cuartil-1, cuartil-2, cuartil-3**  **3) Rango Intercuartilico**  **4) varianza (use formula computacional)**  **5) coeficiente de variación**  **📋 Assignment #2**  **Universidad Politécnica de Puerto Rico**  **ASIGNACION**  **ENGI 2210/2270 Probability and Statistics for Engineers**  **Nombre:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Num:Est:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**  **Prof. Ing. José Raúl Díaz, PE;**    **I. Realicé el siguiente problema:**  **Para los siguientes datos: (Ordenados) ∑X=880 ; ∑X2=39648**  **30 34 37 38 38 40 42 42 42 42**  **44 44 47 48 48 49 50 52 56 57**  **a) Realice Tabla de clases, Intervalos ó fronteras , frecuencias relativa que se requiere para un Dibuje un Diagrama de Frecuencia Relativa. Utilice 6 Clases**  **b) Calcule**  **1) la media o promedio 5) Rango**  **2) quartil-1 6) Varianza**  **3) quartil-2 7) desviación estándar**  **4) quartil-3 8) coeficiente de variación**  **II Determine : El coeficiente de variación (CV) de los dos conjuntos de datos.**   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | **Data1** | **11** | **12** | **12** | **14** | **15** | **20** | | **Data2** | **8** | **8** | **10** | **10** | **15** | **15** | |
| **🕮 REPASO DEL MATERIAL DEL MODULO** |
| **Sera removido de este modulo**      **Extraida de Power point y manipulada por Paint del libro :David Doane and Lori Seward; Applied Statistics in Business and Economics; McGraw-Hill Irwin, NY 2007, USA chapter 4**            **Extraida de Power point y manipulada por Paint del libro :David Doane and Lori Seward; Applied Statistics in Business and Economics; McGraw-Hill Irwin, NY 2007, USA chapter 4**    **Extraida de Power point y manipulada por Paint del libro :David Doane and Lori Seward; Applied Statistics in Business and Economics; McGraw-Hill Irwin, NY 2007, USA chapter 4**      **Extraida de Power point y manipulada por Paint del libro :David Doane and Lori Seward; Applied Statistics in Business and Economics; McGraw-Hill Irwin, NY 2007, USA chapter 4**      **Extraida de Power point y manipulada por Paint del libro :David Doane and Lori Seward; Applied Statistics in Business and Economics; McGraw-Hill Irwin, NY 2007, USA chapter 4**      **Extraida de Power point y manipulada por Paint del libro :David Doane and Lori Seward; Applied Statistics in Business and Economics; McGraw-Hill Irwin, NY 2007, USA chapter 4**      **Extraida de Power point y manipulada por Paint del libro :David Doane and Lori Seward; Applied Statistics in Business and Economics; McGraw-Hill Irwin, NY 2007, USA chapter 4** |
| **🗎 Self Assessment (How to Know if you’re ready for grading)** |
| **Check your understanding:** |
| **🗔 Assessment (How to show me what you’ve learned)**  **Select the assessment tool to used: (Test or SafeAssignment)** |
| **(Include due dates, time frames and point possible).** |
| **⭯ Virtual Group Activity**  **Select the tool to used: (Blog, Discussion Board, Journal, Wiki, Group, Chat or Virtual Classroom)** |
| **(Include due dates, time frames and point possible).** |
| **⏭ SITUACION ESPECIAL EN EL COMPUTO DE CUARTILES POR MEDIOS ELECTRONICOS DE SOFTWARE-PROGRAMAS ESTADISTICOS** |
| **Lamentablemente en la actualidad no se logro un conceso en el computo de los cuartiles mediante medios electrónicos entiéndase mediante hojas electrónicas y programas estadísticos.**  **Ojo- hay problemas en este sentido:**  **Problemas con el computo de cuartiles utilizando software o programas estadísticos.**  **Asunto de computo de cuartiles: 5 10 12 23 24 30   Por analisis visual : Q1=10   Q2=(12+23)/2=17.5     Q3=24  Podras observar que el los "Screen Saving" de varios programas errores en computo de Q1 y Q3  NO hay problemas con Q2=Md=P50=Mediana  a. MINI-TAB    Calcula erróneamente Q1=8.75    y   Q3=25.50  b.Excel            Calcula erróneamente  Q1=10.5   y   Q3 = 23.75  c.Solo un programa: STATDISK  calculo correctamente Q1=10  y Q3=24**  **Moraleja: " No podemos creer ciegamente los resultados de un programa de computadora. Sobre todo si esta reñido con el sentido común y de lógica matemática."**  **Debes validar cuidadosamente el programa para poder utilizarlo correctamente y tomar decisiones**          **DIFERENCIAS EN COMPUTO DE CUARTILES**  **Comparación de programas que calculan cuartiles y sus efectos, ver página del libro de Mario Triola Elementary Statistics. A continuación**  **Quartiles0001.jpg**  **Foto de pagina112 capitulo 3;Triola Mario; Elementary Statistics, Ed 10, 2006 Addison Wesley-Pearson USA** |
| **References**  **David Doane and Lori Seward; Applied Statistics in Business and Economics; McGraw-Hill Irwin, NY 2007, USA**  **Mario Triola; Elementary Statistics ; 10th ed.,2006, Pearson Educational Inc. USA**  **Mark Berenson, David Levine and Timothy Krehbiel; Basic Business Statistics Concepts and Applications; 9th edition, 2004; Pearson Prentice-Hall, USA**  **Ken Black, David Eldredge3; Business & Economics Statistics; Using Microsoft Excel; South-Western; Thomson Learning, 2002; USA**  **William Menderhall, Robert Beaver y Barbara Beave; Introducción a la probabilidad y estadística; 1ra. Edición 2002 International Thomson Editores, México DF**  **Ken Black; Business Statistics Contemporary Decision Making; 3rd edition;2001; South-Western College Publishing Thomson Learning, USA**  **Required Resources**  **Additional Resources** |
| **📋 REFERENCIAS Power Point / Youtube** |
| **PowerPoint- De los autores : David Doane and Lori Seward; Applied Statistics in Business and Economics; McGraw-Hill Irwin, NY 2007, USA**    **YOUTUBE- VIDEO REFERENCES**  **Referencia de página en internet con curso de estadística**  [**www.apstatsguy.com**](http://www.apstatsguy.com)  [**www.mathtutordvd.com**](http://www.mathtutordvd.com)  **http://aula.tareasplus.com/Marcela-Gomez/Probabilidad-y-Estadistica** |