Hay en total 6,435 grupos diferentes de 8 maestros, que se pueden formar.

Por lo cual, la probabilidad de que al seleccionar al azar un grupo de 8 maestros, es formado por 5 maestros y 3 maestras, será el número de casos favorables entre número total de casos posibles, es decir

$$P(5H,3M) = \frac{2,520}{6,435}$$

$$= \frac{56}{143}$$

$$= 0.3916$$

O lo que es lo mismo, la probabilidad es de 39.16%

Ejercicio 3.4

1. ¿Cuántos comités diferentes formados por 2 estudiantes de primer semestre 3 estudiantes de segundo semestre se pueden seleccionar entre 15 estudiantes de primer semestre y 12 estudiantes de segundo semestre?

2. Una urna contiene 3 boletas rojas y 6 boletas azules. ¿De cuántas formas pueden sacar 3 boletas, si dos de ellas deben ser rojas?

3. ¿Cuántas palabras diferentes de dos vocales y dos consonantes se pueden formar con 7 consonantes diferentes y 5 vocales diferentes?

4. El mismo problema 3, si las consonantes y las vocales deben ir alternadas.

5. ¿Cuántas rectas se pueden trazar por 14 puntos coplanares, si tres cualesquiera de ellos, no son colineales?

6. Si se dispone de 10 personas para hacer una guardia nocturna de 3 personas ¿durante cuántas noches se podrá tener una guardia diferente?

7. Una bolsa contiene 6 bolas rojas, 3 bolas amarillas y 4 bolas azules. ¿De cuántas maneras se podrán extraer 8 bolas, de tal forma que sean 3 rojas, amarillas y 3 azules?

8. Con los datos del problema # 7, pero considerar que en las 8 bolas extraídas por lo menos 4 sean rojas.

9. Comprobar que $C(6,2) = C(7,3)-C(6,3)$

10. Si $C(n,r) = 20$ y $P(n,r) = 120$, encontrar los valores que deben tener "n" y "r"



11. En un grupo de la preparatoria # 27 de la U.A.N.L. hay 18 hombres y 12 mujeres, en un segundo grupo hay 14 hombres y 16 mujeres. Si se selecciona un grupo al azar y de él se escoge al azar una persona, ¿cuál es la probabilidad de que la persona sea mujer?

12. Un niño tiene 15 pares de calcetines, 9 de los cuales son blancos y 6 son negros ¿cuál es la probabilidad de que al escoger 8 pares de calcetines, 6 de ellos sean blancos y 2 sean negros?

13. Al tirar una moneda 6 veces, ¿cuál es la probabilidad de que caigan al menos 4 caras?

14. Al tirar una moneda 6 veces, ¿cuál es la probabilidad de que caigan exactamente 4 sellos?

15. Si se tiran 6 monedas simultáneamente, ¿cuál es la probabilidad de obtener exactamente 4 sellos?

### 3.5 Estadística descriptiva y análisis de datos

Objetivo:

| |
|---|
| Dado un conjunto de datos, determinar la distribución: de frecuencias agrupadas, de frecuencias acumuladas y de porcentajes acumulados; graficar:el histograma y el polígono de frecuencias de la distribución; calcular: la media, la mediana, la moda y la desviación estándar de la distribución. Y dada una calificación individual, la media y la desviación estándar de una distribución normal, calcular: su calificación estándar equivalente y su percentil correspondiente. |

Si preguntáramos al "hombre de la calle" qué significa para él la Estadística, en el mejor de los casos obtendríamos como respuesta: "La Estadística es una palabra mágica, es el 'abracadabra', que nos permite manejar los números a nuestro antojo. Manipulando los números de acuerdo con ciertas reglas secretas y muy bien guardadas, podremos probar lo que deseamos". O simplemente: "Estadística es solamente una colección de hechos". Incluso podríamos encontrar alguna persona más conocedora que opinase que es una "colección de hechos y circunstancias pasadas, de interés vital, tales como el número de bañeras existentes en un país en 1929.

Es verdad que éstas y otras actividades se atribuyen frecuentemente al campo de las estadísticas. No es verdad, sin embargo, que los Estadísticos se ocupen de ellas. ¿Qué es entonces la Estadística? Aunque es virtualmente imposible obtener un consenso general de la definición de Estadística, es factible diferenciar dos definiciones de Estadística.

1. La Estadística es comúnmente considerada como una colección de hech numéricos expresados en términos de una relación resumida, y que han s recopilados a través de varias observaciones, o a partir de otros datos numéricos, considerada la Estadística está constituida por una colección de enunciados ta como, "El cociente intelectual promedio de los niños de 5° de primaria es...," o "si de cada diez personas prefieren el licor X cuando se compara con el licor Y", o bi "los campeones mundiales de fútbol anotaron 25 goles en 10 partidos realiza frente a seleccionados de diferentes países..."

2. La Estadística puede igualmente ser considerada como un método para tratar da numéricos. Esta definición refuerza el concepto de que la Estadística es instrumento que se orienta a la recolección, organización, y análisis de da numéricos de observaciones.

La primera función constituye el objeto de esta sección.

Deben diferenciarse dos funciones del método estadístico: técnicas de Estadíst descriptiva y técnicas de Inferencia Estadística o inductiva.

El objetivo central de la Estadística descriptiva es presentar información en for conveniente, útil y comprensible. La inferencia estadística, por su parte se ocupa generalizar esta información, o, más específicamente hace inferencias acerca poblaciones a partir de las muestras extraídas de estas poblaciones.

**Estadística descriptiva**

Cuando un investigador conduce un estudio, de ordinario reúne gran cantidad información numérica o datos acerca del problema en cuestión. Los datos pod tener variedad de formas: datos de frecuencias (recuentos de votantes que prefie uno u otro candidato político), o datos escalares (los pesos netos de los ingredien de un cereal de consumo popular, o los cocientes intelectuales de un grupo estudiantes universitarios). En su forma original, tal cual son recopilados, estos da son usualmente un enredo de calificaciones, recuentos, frecuencias, etc. Al reali la función descriptiva el Estadístico formula reglas y procedimientos para presentación de los datos en una forma más útil y significativa. Así, el Estadíst establece reglas a través de las cuales los datos pueden representarse gráficamen También formula reglas para calcular varios estadísticos a partir de los da originales.

Imaginemos que un instructor consejero quiere someter a un grupo de estudiantes secundaria a una serie de pruebas (para medir su inteligencia, aptitud, personalid etc.). ¿Qué tipo de operaciones podrá ejecutar con las calificaciones y medi resultantes para cumplir con la función descriptiva?



1. Podrá reorganizar las calificaciones y agruparlas de varias formas, para obtener así una visión global de todo el conjunto de datos.

2. Podrá construir tablas, gráficas y figuras que permitan visualizar los resultados.

3. Podrá convertir los resultados originales a formas que sean más útiles para propósitos específicos. Así, podrá convertir estas calificaciones originales en rangos percentiles.

4. Podrá también calcular promedios, para aprender algo acerca del comportamiento específico de sus problemas.

5. Empleando el promedio como punto de referencia podrá describir la dispersión de calificaciones con respecto a un punto central. Las estadísticas que cuantifican esta dispersión son conocidas como medidas de variabilidad o medidas de dispersión.

**Inferencia estadística**

El trabajo de un investigador no ha llegado a su fin cuando ha concluido con la función descriptiva. Por el contrario, está frecuentemente más cerca del principio, que del fin de su tarea. La razón de esta afirmación es obvia cuando consideramos que el propósito de su investigación es a menudo explorar hipótesis de naturaleza general, más que simplemente comparar unas cuantas muestras.

Imaginemos que eres un farmacólogo que está interesado en conocer los efectos de una droga determinada sobre la ejecución de una tarea que implica coordinación psicomotora. En consecuencia, usted diseña un estudio que implique dos condiciones: una experimental y otra de control. Tu suministras la droga a los sujetos en experimentación en períodos de tiempo específicamente determinados antes de empezar la prueba. Para evitar "efectos de placebo" se dará una píldora inocua que contenga ingredientes inertes a los sujetos de control. Después de que todos los sujetos hayan sido sometidos a la prueba, elaborarás la función descriptiva. Así hallarás que "en promedio" los sujetos experimentales no actuaron tan bien como los de control. En otras palabras, la media aritmética de los datos del grupo experimental será más baja que la del grupo de control. En consecuencia, te preguntarás "¿puedo concluir que la droga produjo tal diferencia entre los grupos?" o, una pregunta aún más general, "¿puedo asegurar que la droga tiene un efecto adverso sobre la ejecución de la tarea que se investiga?" Para responder estos interrogantes, no es suficiente basarse solamente en la Estadística Descriptiva.

"Después de todo", razonarás, "aún si la droga no tuviera efecto, es altamente improbable que los promedios de los dos grupos hubiesen sido idénticos. Alguna diferencia habrá de observarse". La intervención de variables incontrolables (algunas veces llamadas en forma imprecisa "factores de azar") ciertamente produce alguna disparidad entre las medidas de los grupos. La pregunta crítica, desde el punto de vista de la inferencia estadística, será: ¿es la diferencia suficientemente grande com para descartar las variaciones incontroladas en el experimento como una explicació suficiente? Enunciándolo de otra manera, si fuésemos a repetir el experimen ¿seríamos capaces de predecir con confianza, que las mismas diferencias (ejempl una medida es más grande que la otra) se repetiría sistemáticamente?

En el momento en que surgen estas preguntas, nos movemos hacia el área fascinan del análisis estadístico conocida como inferencia estadística o Estadística Inductiv

## Distribuciones de frecuencias

Imaginemos que acabas de aceptar el cargo de director de una importante escuela d bachillerato. Tu responsabilidad es preparar una reforma curricular acorde con la necesidades y motivaciones de los estudiantes y que, al mismo tiempo, estimule su capacidades intelectuales. Desde luego, está fuera del alcance del texto solucion este problema tan interesante y complejo. Sin embargo, no se puede avanzar hac una solución sin evaluar antes las capacidades intelectuales del cuerpo estudiantil. E consecuencia, acudes a los archivos y "extraes información" al azar (es decir, de t forma que cada miembro de la población tenga la misma oportunidad de s seleccionado) tomando 110 expedientes de estudiantes, que brinden informaci personal y académica. Puesto que su interés presente es valorar la capacid intelectual, tu enfocarás su atención hacia la información sobre el "C.I. estimado Al anotar estos datos obtendrás las cifras señaladas en la tabla 1

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 154 | 131 | 122 | 100 | 113 | 119 | 121 | 128 | 112 | 93 |
| 133 | 119 | 115 | 117 | 110 | 104 | 125 | 85 | 120 | 135 |
| 116 | 103 | 103 | 121 | 109 | 147 | 103 | 113 | 107 | 98 |
| 128 | 93 | 90 | 105 | 118 | 134 | 89 | 143 | 108 | 142 |
| 85 | 108 | 108 | 136 | 115 | 117 | 110 | 80 | 111 | 127 |
| 100 | 100 | 114 | 123 | 126 | 119 | 122 | 102 | 100 | 106 |
| 105 | 111 | 127 | 108 | 106 | 91 | 123 | 132 | 97 | 110 |
| 150 | 130 | 87 | 89 | 108 | 137 | 124 | 96 | 111 | 101 |
| 118 | 104 | 127 | 94 | 115 | 101 | 125 | 129 | 131 | 110 |
| 97 | 135 | 108 | 139 | 133 | 107 | 115 | 83 | 109 | 116 |
| 110 | 113 | 112 | 82 | 114 | 112 | 113 | 142 | 145 | 123 |

Tabla 1
**Resultados del C.I. de 110 estudiantes de bachillerato seleccionados al azar.**

Como es obvio, en principio estas cantidades no tienen para ti "pies ni cabeza" menos que las organices de un modo sistemático. Quizás puedas idear ordenamiento de todos los resultados en forma descendente y luego colocar una rayi vertical a su lado cada vez que ocurran. (tabla 2).



<table>
<tr><th>X</th><th>f</th><th>X</th><th>f</th><th>X</th><th>f</th><th>X</th><th>f</th></tr>
<tr><td>154</td><td>|</td><td>135</td><td>||</td><td>116</td><td>||</td><td>97</td><td>||</td></tr>
<tr><td>153</td><td></td><td>134</td><td>|</td><td>115</td><td>||||</td><td>96</td><td>|</td></tr>
<tr><td>152</td><td></td><td>133</td><td>||</td><td>114</td><td>||</td><td>95</td><td></td></tr>
<tr><td>151</td><td></td><td>132</td><td>|</td><td>113</td><td>||||</td><td>94</td><td>|</td></tr>
<tr><td>150</td><td>|</td><td>131</td><td>||</td><td>112</td><td>|||</td><td>93</td><td>||</td></tr>
<tr><td>149</td><td></td><td>130</td><td>|</td><td>111</td><td>|||</td><td>92</td><td></td></tr>
<tr><td>148</td><td></td><td>129</td><td>|</td><td>110</td><td>|||||</td><td>91</td><td>|</td></tr>
<tr><td>147</td><td>|</td><td>128</td><td>||</td><td>109</td><td>||</td><td>90</td><td>|</td></tr>
<tr><td>146</td><td></td><td>127</td><td>|||</td><td>108</td><td>||||||</td><td>89</td><td>||</td></tr>
<tr><td>145</td><td>|</td><td>126</td><td>|</td><td>107</td><td>||</td><td>88</td><td></td></tr>
<tr><td>144</td><td></td><td>125</td><td>||</td><td>106</td><td>||</td><td>87</td><td>|</td></tr>
<tr><td>143</td><td>|</td><td>124</td><td>|</td><td>105</td><td>||</td><td>86</td><td></td></tr>
<tr><td>142</td><td>||</td><td>123</td><td>|||</td><td>104</td><td>||</td><td>85</td><td>||</td></tr>
<tr><td>141</td><td></td><td>122</td><td>||</td><td>103</td><td>|||</td><td>84</td><td></td></tr>
<tr><td>140</td><td></td><td>121</td><td>||</td><td>102</td><td>|</td><td>83</td><td>|</td></tr>
<tr><td>139</td><td>|</td><td>120</td><td>|</td><td>101</td><td>||</td><td>82</td><td>|</td></tr>
<tr><td>138</td><td></td><td>119</td><td>|||</td><td>100</td><td>||||</td><td>81</td><td></td></tr>
<tr><td>137</td><td>|</td><td>118</td><td>||</td><td>99</td><td></td><td>80</td><td>|</td></tr>
<tr><td>136</td><td>|</td><td>117</td><td>||</td><td>98</td><td>|</td><td></td><td></td></tr>
</table>

Tabla 2
**Distribución de frecuencias de los resultados del C.I. de 110 estudiantes de bachillerato seleccionados al azar.**

El número de rayitas representa pues la frecuencia con que aparece cada resultado.

Al terminar se habrá construido un conjunto no agrupado de distribución frecuencias de resultados. Advierta que, en el presente ejemplo, algunos punta tienen una extensión muy vasta, mientras que otros presentan una frecuencia de c y no hay "visualización" clara que indique una tendencia central. En estas condicio la mayoría de los investigadores acostumbran "agrupar" los puntajes en lo que llama intervalos de clase, para obtener una distribución de frecuencias de "punta agrupados".

El agrupamiento en intervalos de clase implica una especie de "ruptura de la esca en la cual asignamos los puntajes mutuamente excluyentes[1] en donde éstas definen en términos de los intervalos de grupo empleados. Las razones para ha este tipo de agrupación son dobles:

(1) es antieconómico y poco práctico tratar con un gran número de cas distribuidos en muchos puntajes a menos que dispongamos de calculado automáticas.

(2) Algunos de los puntajes tienen asociada una frecuencia tan baja que no justifica mantenerlos como entidades distintas y separadas.

Como factor negativo podemos mencionar que al agrupar los puntajes se pie inevitablemente parte de la información. Por ejemplo, las calificaciones individu no se identifican cuando haya un agrupamiento en intervalos de clase y son inevitab pequeños errores estadísticos basados en su agrupación.

La próxima pregunta es "¿Qué base emplearemos para agrupar los datos en intervalos de clase que vamos a utilizar?" Obviamente, el intervalo seleccionado deberá ser tan amplio que perdamos la discriminación proporcionada por nues medida original. Por ejemplo, si tuviéramos que dividir en dos clases los puntajes C.I. ya seleccionado, y lo hiciéramos entre los inferiores a 100, y los iguale superiores a 100; en la práctica toda la información inherente a los puntajes origina se habría perdido. Por otra parte, los intervalos de clase no deben ser tan peque que se desvirtúe el objetivo que se busca con la agrupación. Para responder a nues pregunta, no hay desgraciadamente ninguna norma general que pueda ser aplicada todos los casos. La mayoría de las veces el escoger el número de intervalos de cl debe representar un juicio basado en la influencia que tienen los efectos d agrupación sobre la discriminación de los datos y sobre la economía de presentación. No obstante, se acepta generalmente que buena parte de los da obtenidos en las ciencias que estudian el compartimiento puede ser agrupada en a 20 intervalos de clase. (Consideraremos 15 intervalos de clase para los datos se analizan en este ejemplo).

Habiendo decidido el número apropiado de intervalos de clase para una serie de da son relativamente sencillos los procedimientos para asignar resultados a los interva de clase. Como también es posible utilizar diferentes técnicas, nosotros emplearem

[1] Llamamos a las clases mutuamente excluyentes porque es imposible para un puntaje de una m pertenecer a más de una clase.



sólo una para ser siempre consecuentes. El procedimiento que debe emplearse es el siguiente:

1er Paso. Encuentra la diferencia entre el resultado más alto y el más bajo en los datos originales. Añade 1 para obtener el número total de resultados o de resultados potenciales. En el presente ejemplo, el resultado es (154-80) + 1 = 75.

2do Paso. Divide esta cifra por 15 para obtener el número de resultados o de resultados potenciales en cada intervalo. Si el valor resultante no es un número entero, y generalmente no lo es, los autores prefieren redondear al número entero más cercano, esta práctica está lejos de ser universal, y no se comete ningún error si se redondea al número más cercano. En este ejemplo, el número de calificaciones para cada intervalo de clase es 75/15, o sea 5. Identificaremos este número con el símbolo *i*. En el ejemplo, *i* = 5.

3er Paso. Toma el resultado más bajo de los datos originales como el límite inferior del primer intervalo de clases. Agrega *i*-1 para obtener el puntaje máximo del primer intervalo de clase. En esta forma, el intervalo de clase mínimo en los datos que tenemos a mano es 80-84.

4t Paso. El límite inferior de intervalo de clase siguiente, será el número entero consecutivo al máximo puntaje del intervalo de clase inferior. En el presente ejemplo, el próximo número entero es 85. Sigue las mismas etapas del 3er paso para obtener el puntaje más alto del segundo intervalo de clase. Prosigue con este procedimiento para cada intervalo de clase superior sucesivo, hasta que todos los puntajes se hayan incluido en sus intervalos de clase apropiados.

5t Paso. Asigna cada puntaje obtenido al intervalo de clase dentro del cual está incluido. La distribución de frecuencias agrupada que aparece en la tabla 3 se obtuvo empleando el procedimiento ya citado.

| INTERVALO DE CLASE | f | INTERVALO DE CLASE | f |
|---|---|---|---|
| 150-154 | 2 | 110-114 | 17 |
| 145-149 | 2 | 105-109 | 14 |
| 140-144 | 3 | 100-104 | 12 |
| 135-139 | 5 | 95-99 | 4 |
| 130-134 | 7 | 90-94 | 5 |
| 125-129 | 9 | 85-89 | 5 |
| 120-124 | 9 | 80-84 | 3 |
| 115-119 | 13 | | |
| N = 110 | | | |

Tabla 3
**Distribución de frecuencias agrupada de los resultados del C.I. basadas en los datos que aparecen en la tabla 2.**