# MA34B – Estadística
# Estimación Puntual

Prof. Rodrigo Abt B.
rabt@dim.uchile.cl

# Problema

- El gerente de un supermercado ha encargado el estudio de una ampliación del sitio destinado a estacionamientos.
- Si la probabilidad de tener más de 100 autos en un día es mayor al 70% se deberá hacer la ampliación.
- De estudios similares, se sabe que el número de autos que llega en un día al estacionamiento sigue una distribución de Poisson.
- Se toma una muestra de 5 días en los cuales se mide la cantidad de autos que llegó al estacionamiento, y se obtienen los valores: 99, 102, 112, 97,130.

# Preguntas

*¿Cómo calculo la probabilidad?*

# Planteamiento Formal

- Claramente para el cálculo de la probabilidad necesitamos conocer la tasa de la distribución de Poisson.
- Podemos escribir el problema anterior de manera más formal:
  - Sea $X_1,X_2,\ldots,X_n$ una m.a.s. de una v.a. X que sigue una distribución de Poisson con tasa $\lambda$.
  - ¿Cuánto vale $\lambda$?

# Estimadores y Estimaciones (1)

- Una manera natural de dar un valor para el parámetro buscado es con el promedio de los datos:

$$\hat{\lambda} = \frac{1}{5}(99 + 102 + 112 + 97 + 130) = 108$$

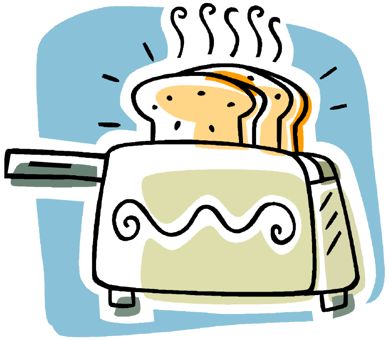
- Este valor corresponde a una *estimación* del parámetro.
- Mientras que la fórmula:

$$\hat{\lambda} = \frac{X_1 + X_2 + X_3 + X_4 + X_5}{5}$$

corresponde a un *estimador*.

# Estimadores y Estimaciones (2)

- Un estimador es una fórmula, una expresión, que permite dar valores a parámetros desconocidos.
- Siguiendo la analogía de una preparación culinaria, el estimador correspondería a la receta, la muestra aleatoria de valores a los ingredientes, y la comida preparada a la estimación.



# Características deseables

- Pero, en el problema planteado, ¿es realmente $\lambda$=108 una buena estimación?, o mejor dicho, es el promedio un buen estimador?
- ¿Qué entendemos por “bueno”?
- Un buen estimador debería al menos en promedio “achuntarle” al verdadero valor del parámetro estimado, es decir, debería ser tal que si $\hat{\lambda}$ es un estimador de $\lambda$, Es decir:

$$E(\hat{\lambda}) = \lambda$$

- Si un estimador cumple con lo anterior se le denomina *insesgado*, y a la cantidad $E(\hat{\lambda}) - \lambda$ se le denomina “sesgo”

# Precisión

- Sabemos que el valor de una estimación puede variar según la muestra que se tome, y por ende debemos tratar de evitar que las estimaciones sean muy dispersas.
- Debemos encontrar idealmente un estimador cuya precisión se pueda determinar para varias muestras, y además, que aumente a medida que crece el tamaño de la muestra.
- En este sentido, la varianza del estimador representa el grado de precisión del mismo.
- Entonces, preferiremos aquellos estimadores cuya varianza disminuya al crecer el tamaño de muestra “n”

# Métodos Para Estimar

- ¿Cómo encontramos un estimador?
- Respuesta: a través de métodos de estimación. Los más comunes son:
  - Método de los Momentos
  - Método de Máxima Verosimilitud
  - Método de Bayes
- NOTA: El último método requiere de supuestos adicionales.

# Método de los Momentos

- La ley de los grandes números nos muestra que

$$m_r = \frac{1}{n}\sum X_i^r \to \mu_r \text{ (c.s.)}$$

- Luego, sería natural estimar el momento de orden teórico con el momento de orden muestral, es decir:

$$\widehat{\mu}_r = m_r$$

# Método de Máxima Verosimilitud

**Problema**: Supongamos que se escoge una m.a.s. de 8 pacientes en un hospital para determinar la tasa de contagio de una cierta infección en todo el lugar.

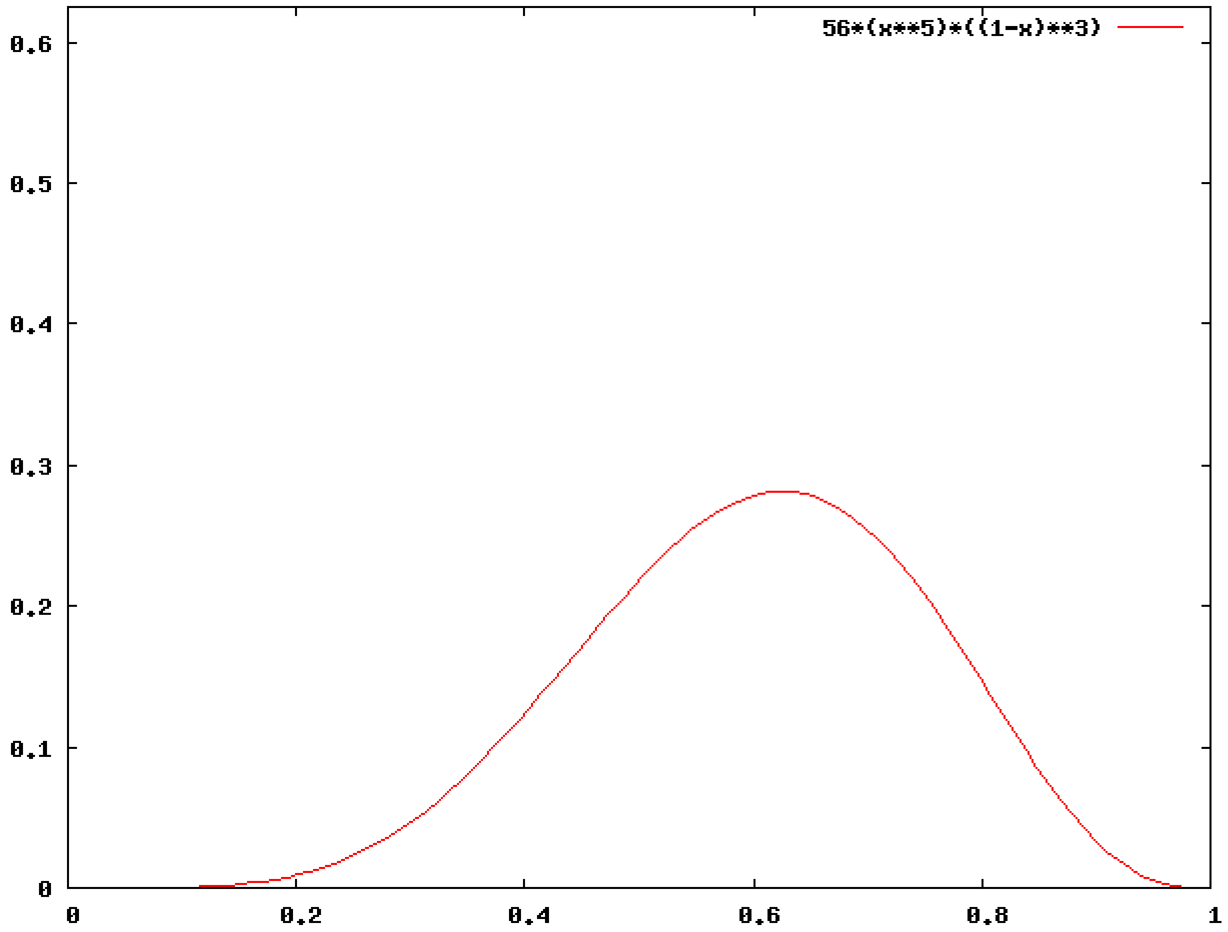
Sean

x1=1,x2=0,x3=0,x4=0,x5=0,x6=1,x7=1,x8=1

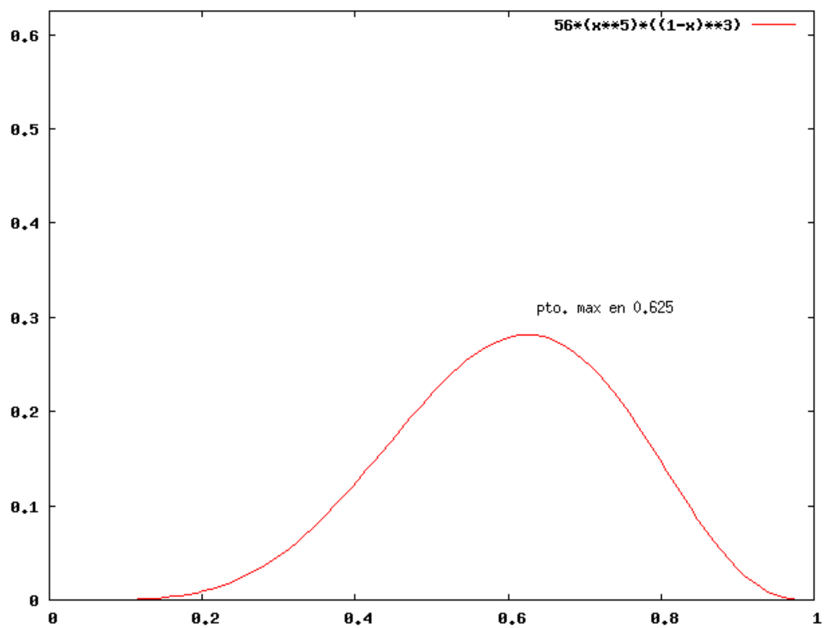
los resultados de los exámenes a los pacientes de la muestra, siendo 1=contagio,0=sin contagio. Este experimento se puede modelar (sin perder generalidad) a través de una distribución Binomial(8,p). Sea

$$W(p) = \Pr(Z=5) = \binom{8}{5} p^5 (1-p)^3$$

# Gráfica del problema



# Solución




Podemos observar que el valor de *p* que hace más probable obtener la muestra que obtuvimos es 0.625, que correspoonde al máximo de la función W(p)

# El Estimador Máximo Verosímil

- Se llama función de verosimilitud a la densidad conjunta del vector conformado por los valores muestrales $x_1, x_2, \ldots, x_n$:

$$f(x_1, x_2, \ldots, x_n/\theta) = \prod_{i=1}^{n} f(x_i/\theta)$$

- El estimador máximo verosímil es aquella expresión $\theta = \theta(x_1, x_2, \ldots, x_n)$ que hace máxima la función de verosimilitud.

# Distribución a Posteriori

- El método de Bayes asume la existencia de una distribución a priori para el parámetro desconocido, por lo cual se convierte en v.a. a su vez.
- En este caso se procede a recalcular la distribución del parámetro, dados los resultados de la muestra, obteniéndose una distribución a posteriori:

$$\xi(\theta|x) = \frac{f_n(x/\theta)\pi(\theta)}{\int_{\Theta} f_n(x/\theta)\pi(\theta)d\theta}$$

# Estimadores de Bayes

- Si consideramos una función de pérdida L(θ,δ(x)), en que δ(x) es un estimador de θ, entonces es natural encontrar un estimador que minimice dicha pérdida.
- Un estimador de Bayes es solución de:

$$min_\delta \quad E[L(\theta, \delta(x)]$$

- Existen diversas funciones de pérdidas, como la cuadrática, la absoluta y la pérdida "0-1", siendo la más utilizada, la cuadrática.
- Para el último caso, el estimador de Bayes es simplemente el valor esperado de θ|X, o sea, se obtiene con la esperanza de la distribución a posteriori.