

Vídeos docentes sobre
Probabilidad y Teoría de la Decisión

Primer ejemplo: planteamiento del problema

F. J. Díez Vegas
Dpto. Inteligencia Artificial. UNED

fjdiez@dia.uned.es
www.ia.uned.es/~fjdiez

Primer ejemplo: distribución bivaluada

- ◆ Enfermedad: tasa de supervivencia, 60%
- ◆ Objetivo: evaluar la efectividad del XYZ, un nuevo fármaco
- ◆ Tres estudios realizados sobre el XYZ
 - Hospital A: 20 pacientes, 15 supervivientes (75%)
 - Hospital B: 20 pacientes, 17 supervivientes (85%)
 - Hospital C: 60 pacientes, 45 supervivientes (75%)
- ◆ Preguntas:
 - ¿Es efectivo el medicamento
o se deben al azar estos resultados favorables?
 - ¿Qué estudio aporta mayor evidencia a favor del XYZ?
- ◆ Hay que considerar dos factores
 - Tasa de supervivencia observada
 - Tamaño de la muestra

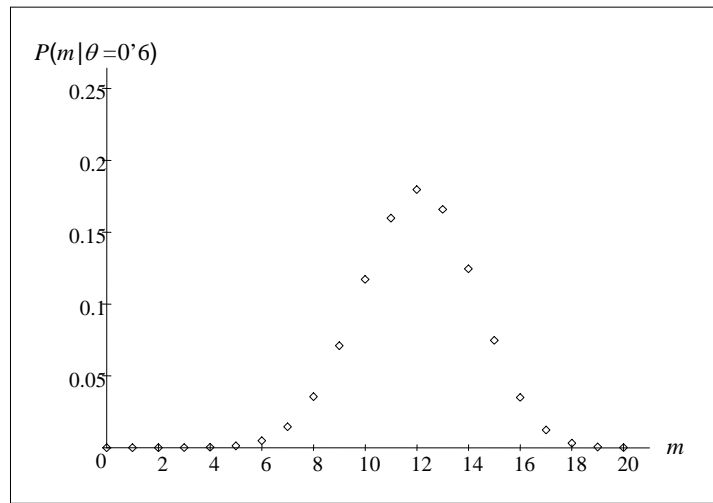
Análisis del problema

- ◆ Dos resultados posibles: fallecimiento, supervivencia
- ◆ Distribución poblacional: bivaluada
 - $P(\text{supervivencia}) = \theta$
 - $P(\text{fallecimiento}) = 1 - \theta$
- ◆ Distribución muestral para el estadístico M (nº supervivientes):

$$P(m|\theta) = \frac{n!}{m!(n-m)!} \theta^m (1-\theta)^{n-m}$$

- ◆ $P(m|\theta)$ es una función de m y de θ
 - Para cada valor del parámetro θ tenemos una función de probabilidad para m
 - Para cada valor m (nº de supervivientes en un experimento) tenemos una función de verosimilitud para θ

$P(m \theta)$		m	m/n	$\theta=0'0$	$\theta=0'2$	$\theta=0'4$	$\theta=0'6$	$\theta=0'8$	$\theta=1'0$
$n = 20$	0	0'00	1	0'012	0'000	0'000	0'000	0'000	0
	1	0'05	0	0'058	0'000	0'000	0'000	0'000	0
	2	0'10	0	0'137	0'003	0'000	0'000	0'000	0
	3	0'15	0	0'205	0'012	0'000	0'000	0'000	0
	4	0'20	0	0'218	0'035	0'000	0'000	0'000	0
	5	0'25	0	0'175	0'075	0'001	0'000	0'000	0
	6	0'30	0	0'109	0'124	0'005	0'000	0'000	0
	7	0'35	0	0'055	0'166	0'015	0'000	0'000	0
	8	0'40	0	0'022	0'180	0'035	0'000	0'000	0
	9	0'45	0	0'007	0'160	0'071	0'000	0'000	0
	10	0'50	0	0'002	0'117	0'117	0'002	0'000	0
	11	0'55	0	0'000	0'071	0'160	0'007	0'000	0
	12	0'60	0	0'000	0'035	0'180	0'022	0'000	0
	13	0'65	0	0'000	0'015	0'166	0'055	0'000	0
	14	0'70	0	0'000	0'005	0'124	0'109	0'000	0
	15	0'75	0	0'000	0'001	0'075	0'175	0'000	0
	16	0'80	0	0'000	0'000	0'035	0'218	0'000	0
	17	0'85	0	0'000	0'000	0'012	0'205	0'000	0
	18	0'90	0	0'000	0'000	0'003	0'137	0'000	0
	19	0'95	0	0'000	0'000	0'000	0'058	0'000	0
	20	1'00	0	0'000	0'000	0'000	0'012	0'000	1

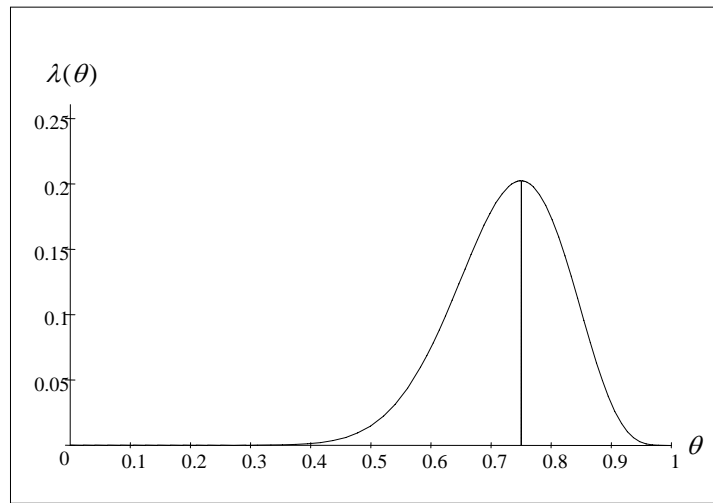


Probabilidad $P(m|\theta=0.6)$ para $n=20$ (Hospitales A y B)

$P(m|\theta)$

$n = 20$

m	m/n	$\theta=0.0$	$\theta=0.2$	$\theta=0.4$	$\theta=0.6$	$\theta=0.8$	$\theta=1.0$
0	0.00	1	0.012	0.000	0.000	0.000	0
1	0.05	0	0.058	0.000	0.000	0.000	0
2	0.10	0	0.137	0.003	0.000	0.000	0
3	0.15	0	0.205	0.012	0.000	0.000	0
4	0.20	0	0.218	0.035	0.000	0.000	0
5	0.25	0	0.175	0.075	0.001	0.000	0
6	0.30	0	0.109	0.124	0.005	0.000	0
7	0.35	0	0.055	0.166	0.015	0.000	0
8	0.40	0	0.022	0.180	0.035	0.000	0
9	0.45	0	0.007	0.160	0.071	0.000	0
10	0.50	0	0.002	0.117	0.117	0.002	0
11	0.55	0	0.000	0.071	0.160	0.007	0
12	0.60	0	0.000	0.035	0.180	0.022	0
13	0.65	0	0.000	0.015	0.166	0.055	0
14	0.70	0	0.000	0.005	0.124	0.109	0
15	0.75	0	0.000	0.001	0.075	0.175	0
16	0.80	0	0.000	0.000	0.035	0.218	0
17	0.85	0	0.000	0.000	0.012	0.205	0
18	0.90	0	0.000	0.000	0.003	0.137	0
19	0.95	0	0.000	0.000	0.000	0.058	0
20	1.00	0	0.000	0.000	0.000	0.012	1

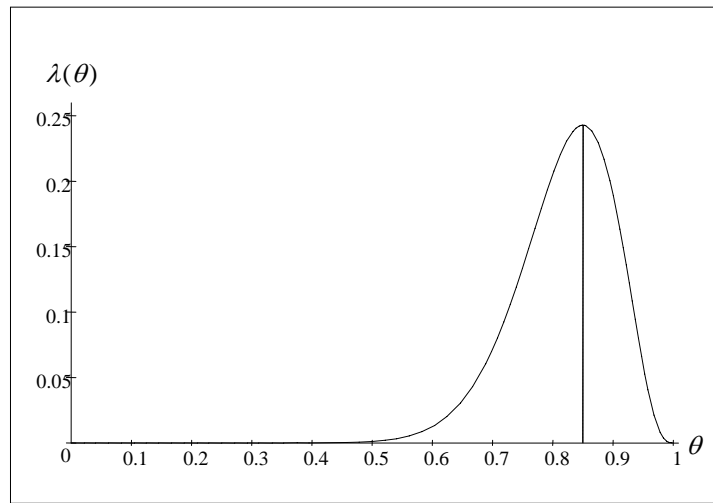


Verosimilitud $\lambda(\theta) = P(m=15|\theta)$ para $n=20$ (Hospital A)

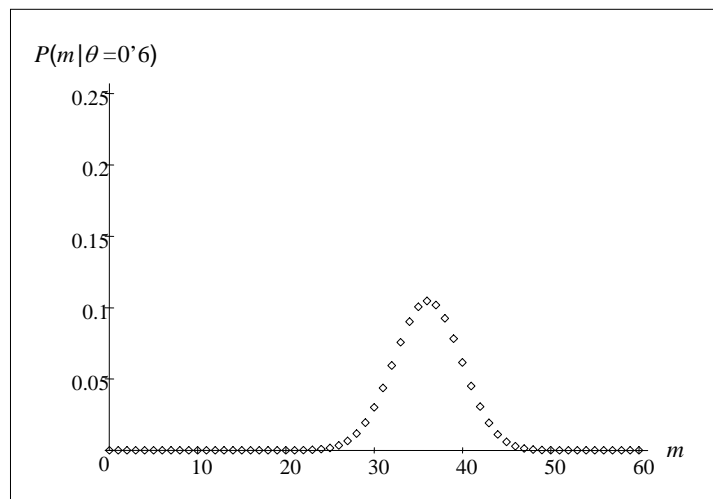
$P(m|\theta)$

$n = 20$

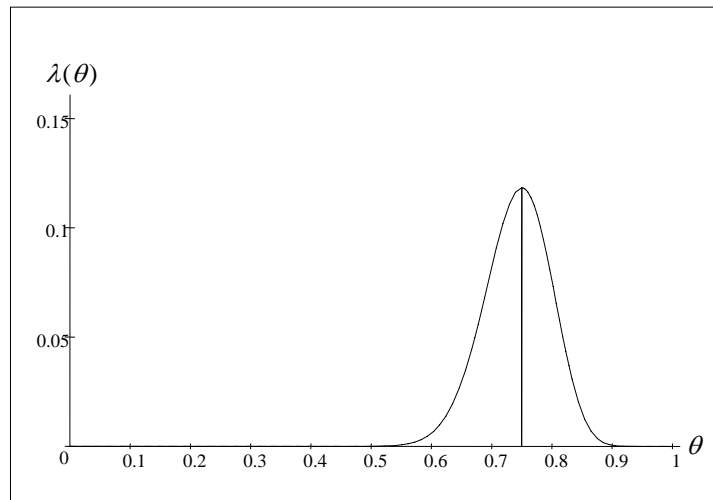
m	m/n	$\theta=0'0$	$\theta=0'2$	$\theta=0'4$	$\theta=0'6$	$\theta=0'8$	$\theta=1'0$
0	0'00	1	0'012	0'000	0'000	0'000	0
1	0'05	0	0'058	0'000	0'000	0'000	0
2	0'10	0	0'137	0'003	0'000	0'000	0
3	0'15	0	0'205	0'012	0'000	0'000	0
4	0'20	0	0'218	0'035	0'000	0'000	0
5	0'25	0	0'175	0'075	0'001	0'000	0
6	0'30	0	0'109	0'124	0'005	0'000	0
7	0'35	0	0'055	0'166	0'015	0'000	0
8	0'40	0	0'022	0'180	0'035	0'000	0
9	0'45	0	0'007	0'160	0'071	0'000	0
10	0'50	0	0'002	0'117	0'117	0'002	0
11	0'55	0	0'000	0'071	0'160	0'007	0
12	0'60	0	0'000	0'035	0'180	0'022	0
13	0'65	0	0'000	0'015	0'166	0'055	0
14	0'70	0	0'000	0'005	0'124	0'109	0
15	0'75	0	0'000	0'001	0'075	0'175	0
16	0'80	0	0'000	0'000	0'035	0'218	0
17	0'85	0	0'000	0'000	0'012	0'205	0
18	0'90	0	0'000	0'000	0'003	0'137	0
19	0'95	0	0'000	0'000	0'000	0'058	0
20	1'00	0	0'000	0'000	0'000	0'012	1



Verosimilitud $\lambda(\theta) = P(m=17|\theta)$ para $n=20$ (Hospital B)



Probabilidad $P(m|\theta=0.6)$ para $n=60$ (Hospital C)



Verosimilitud $\lambda(\theta) = P(m=45 | \theta)$ para $n=60$ (Hospital C)

Tres tipos de inferencia estadística

◆ Estimación puntual

- El valor más verosímil para el Hospital A es $\theta = 0'75$
- El valor más verosímil para el Hospital B es $\theta = 0'85$
- El valor más verosímil para el Hospital C es $\theta = 0'75$

Inconveniente: no nos informa sobre la precisión
(no tiene en cuenta el tamaño de la muestra)

◆ Contraste de hipótesis

- Trata de determinar si el fármaco aumenta la supervivencia
- pero no trata de determinar cuánto la aumenta

◆ Intervalos de confianza

- Trata de determinar la tasa de supervivencia asociada al fármaco