Redes Bayesianas: Test de independencia condicionada

Aritz Pérez[1] Borja Calvo[2]

Basque Center for Applied Mathematics

UPV/EHU

Donostia, Febrero de 2015

Bibliografía

Pardo98: L. Pardo (1998). Teoría de la Información Estadística. Hespérides.

Introducción

- Generalización del test de independencia
- Orígenes en el análisis de tablas de contingencia [Kullback59]
- Basado en propiedades asintóticas de la log. verosimilitud: **información mutua condicionada**

Información mutua condicionada

$$\begin{aligned}
\hat{I}(\mathbf{X}_A;\mathbf{X}_B|\mathbf{X}_C) &= \sum_{\mathbf{x}_C} \hat{p}(\mathbf{x}_C) D_{KL}(\hat{p}(\mathbf{X}_{A,B}|\mathbf{x}_C); \hat{p}(\mathbf{X}_A|\mathbf{x}_C)\hat{p}(\mathbf{X}_B|\mathbf{x}_C)) \\
&= \sum_{\mathbf{x}_{A,B,C}} \hat{p}(\mathbf{x}_{A,B,C}) \log \frac{\hat{p}(\mathbf{x}_{A,B}|\mathbf{x}_C)}{\hat{p}(\mathbf{x}_A|\mathbf{x}_C)\hat{p}(\mathbf{x}_B|\mathbf{x}_C)}
\end{aligned}$$

- Mide la **divergencia de Kullback-Leibler** media de $\hat{p}(\mathbf{X}_A|\mathbf{x}_C) \cdot \hat{p}(\mathbf{X}_B|\mathbf{x}_C)$ respecto a $\hat{p}(\mathbf{X}_{A,B}|\mathbf{x}_C)$
- Determina cuanto se parecen ambas distribuciones que difieren en la dependencia $d(A;B|C)$
- Medida de la **fuerza de la dependencia** $d(A;B|C)$ en los datos

Test de hipótesis

- Establece una **hipótesis nula** H_0
- Herramienta estadística empleada que permite rechazar la hipótesis nula cuando ante evidencias suficientes
- Basado en un estadístico S
- El **estadístico** es una función de los datos, $S(\mathcal{D})$: es una variable aleatoria
- Conocemos la densidad (asintótica) del estadístico bajo la hipótesis nula $f(S|H_0)$
- Conforme más alto es el valor del estadístico la hipótesis nula se vuelve más inverosímil
- **Significatividad** del test α: probabilidad de rechazar la hipótesis nula cuando es verdadera
- **p-valor asociado** a $\hat{s}$: $pval(\hat{s}) = \int_{\hat{s}}^{\text{inf}} f(s|H_0)ds$
- **Rechazamos** la hipótesis nula cuando el $pval(\hat{s}) < \alpha$

Test de independencia condicionada

Hipótesis nula

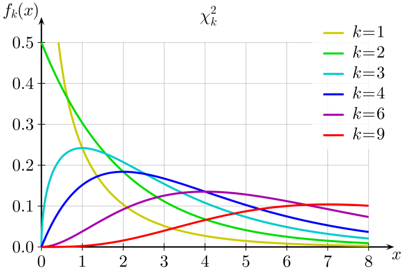
$$H_0 : p(\mathbf{X}_A|\mathbf{x}_C) \cdot p(\mathbf{X}_B|\mathbf{x}_C) = p(\mathbf{X}_{A,B}|\mathbf{x}_C)$$

- La **hipótesis nula** asume la independencia $i(A; B|C)$
- La **significatividad** α cuantifica el **riesgo** que asumimos cuando rechazamos la independencia
- El **pvalor** cuantifica la **credibilidad** de la independencia $i(A; B|C)$

Test de independencia condicionada

Estadístico

$$S(\mathcal{D}) = 2N \cdot \hat{I}(\mathbf{X}_A; \mathbf{X}_B | \mathbf{X}_C) \sim f(S|H_0) \equiv \chi^2_r$$



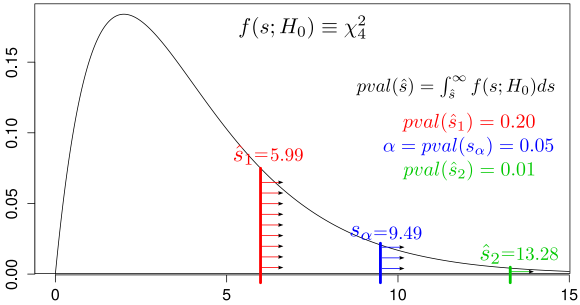
Test de independencia condicionada

- El **estadístico** $S(\mathcal{D})$ se distribuye conforme a una χ^2_r con $r = (r_A - 1)(r_B - 1)r_C$ **grados de libertad**
- El **pvalor** es una medida de la **verosimilitud** de la independencia $i(A; B|C)$
- $S(\mathcal{D})$ **aumenta linealmente** con $\hat{I}(A; B|C)$ y N
- Fijado r, **pvalor disminuye** con $S(\mathcal{D})$
- Fijado $S(\mathcal{D})$, **pvalor aumenta** con r

Procedimiento

1. Queremos decidir si **modelar** la dependencia $d(A; B|C)$
2. Fijamos un nivel de significatividad α: riesgo que estamos dispuestos a asumir
3. **Calculamos** $S(\mathcal{D}) = 2N \cdot \hat{I}(\mathbf{X}_A; \mathbf{X}_B|\mathbf{X}_C)$
4. Consultamos el **p-valor** asociado a S en la distribucion χ^2 con $(|\Omega_A| - 1)(|\Omega_B| - 1)|\Omega_C|$ grados de libertad
5. Si $pval < \alpha$ rechazamos la hipotesis nula $i(A; B|C)$ y asumimos la dependencia $d(A; B|C)$

Resumen



- La independencia condicional se rechaza para el estadístico $\hat{s}_2$ pero no para $\hat{s}_1$