Predicción estadística: conceptos preliminares

Antonio Sala

Dept. Ing. Sistemas y Automatica (DISA)
Universitat Politècnica de València (UPV)

Video-presentación disponible en:

http://personales.upv.es/asala/YT/V/**pre1**.html

Presentación

Motivación:

Muchos cálculos estadísticos y tomas de datos se usan para hacer "predicciones" sobre variables "inciertas".

Objetivos:

Comprender el concepto de "predicción" en estadística (varias interpretaciones).

Contenidos:

Revisión de variables aleatorias. Predicción completa. Predicción puntual; interpretación. Conclusiones.

Variables aleatorias

Variable aleatoria: Posible salida de un experimento, valor incierto de algo.

- Toma valores en un conjunto Ω (finito –cara/cruz–, números reales –dist. normal–)
- Concepto de función de densidad $f : \Omega \rightarrow \mathbb{R}^+$
- Parámetros de una distribución de probabilidad (números acerca de una VA): conceptos de media, varianza

Concepto de predicción

- **Predicción de una variable y:** diferentes opciones.

Supongamos $\Omega \equiv \mathbb{R}$ (variable aleatoria numérica real, usada en la mayor parte de aplicaciones de control).

Concepto "ideal" de predicción:

- Calcular con un modelo o estimar (a partir de muestras) la **función completa de densidad** $f(y)$.
  - Es lo más informativo. Permite, con esa "predicción" calcular la probabilidad de cualquier evento de interés, o cualquier parámetro (media, varianza,...) deseado.
  - Caso más sencillo: si y tiene distribución normal, estimar media y varianza es suficiente para definir la distribución completa.

Concepto de predicción puntual óptima

- **Predicción de una variable** y: diferentes opciones.

1. la **función completa de densidad** $f(y)$.
2. Predicción "**puntual**" p: calcular (o estimar de muestras) un número $p \in \mathbb{R}$ que sea de utilidad...
   - Valor **más probable** (moda, $p = \arg\max f(y)$).
   - Valor **medio** esperado $E[y] = \int y f(y) dy$.

Concepto de predicción puntual óptima

- **Predicción de una variable** y: diferentes opciones.

1. la **función completa de densidad** $f(y)$.
2. Predicción "**puntual**" p: calcular (o estimar de muestras) un número $p \in \mathbb{R}$ que sea de utilidad...
   - Valor **más probable** (moda, $p = \arg\max f(y)$).
   - Valor **medio** esperado $E[y] = \int y f(y) dy$.
   - Predicción mínimo-cuadrática $p = \arg\min_\xi E[(y-\xi)^2]$.
     - Mínima varianza del error de predicción.
     - Coincide con la media $E[y]$: $0 = \frac{dp}{d\xi} = -2E[(y-\xi)] = -2(E[y]-\xi)$
     - Con densidad simétrica alrededor de la media, también coincide con el valor más probable.

Concepto de predicción puntual óptima

- **Predicción de una variable** y: diferentes opciones.

1. la **función completa de densidad** $f(y)$.
2. Predicción "**puntual**" p: calcular (o estimar de muestras) un número $p \in \mathbb{R}$ que sea de utilidad...
   - Valor **más probable** (moda, $p = \arg\max f(y)$).
   - Valor **medio** esperado $E[y] = \int yf(y)dy$.
   - Predicción mínimo-cuadrática $p = \arg\min_\xi E[(y-\xi)^2]$.
     - Mínima varianza del error de predicción.
     - Coincide con la media $E[y]$: $0 = \frac{dp}{d\xi} = -2E[(y-\xi)] = -2(E[y]-\xi)$
     - Con densidad simétrica alrededor de la media, también coincide con el valor más probable.
   - Predicción como "apuesta" con *coste asimétrico* $p = \arg\min_\xi E[J(y-\xi)]$
     - Teoría de decisión: "utilidad", "regret" (coste de error), etc.

Concepto de predicción puntual óptima

- **Predicción de una variable** y: diferentes opciones.
  1. la **función completa de densidad** $f(y)$.
  2. Predicción "**puntual**" p: calcular (o estimar de muestras) un número $p \in \mathbb{R}$ que sea de utilidad...
     - Valor **más probable** (moda, $p = \arg\max f(y)$).
     - Valor **medio** esperado $E[y] = \int yf(y)dy$.
     - Predicción mínimo-cuadrática $p = \arg\min_\xi E[(y-\xi)^2]$.
       - Mínima varianza del error de predicción.
       - Coincide con la media $E[y]$: $0 = \frac{dp}{d\xi} = -2E[(y-\xi)] = -2(E[y]-\xi)$
       - Con densidad simétrica alrededor de la media, también coincide con el valor más probable.
     - Predicción como "apuesta" con *coste asimétrico* $p = \arg\min_\xi E[J(y-\xi)]$
       - Teoría de decisión: "utilidad", "regret" (coste de error), etc.

Discusión

Predicción puntual es más "**sencilla de entender**" (*la media de las pruebas de acceso a la universidad fue de 8.67*), pero **menos informativa** que la función de densidad o histograma (*¿Qué porcentaje de los alumnos sacó más de 7?*).

Un solo número es "poco": la opción intermedia más razonable en aplicaciones es la media + **intervalo de confianza** al α%:

- *la media de las pruebas de acceso a la universidad fue de 8.6, y el 90% de los alumnos tuvo una nota entre 5.9 y 9.7.*
- *Existe un 95% de probabilidad de que mañana llueva entre 10 y 30 l/m^2.*

Conclusiones

- El concepto ideal de predicción implicaría calcular o estimar la función de densidad de una variable.
- En muchos casos, es "complicado" decidir ante tanta información, y se toman decisiones a partir de un único número asociado a dicha variable, en concreto en una "**predicción puntual**", usualmente **moda** o **media**.
- La media mas los intervalos de confianza 95% o 99% suele ser una buena opción "intermedia".

Si la predicción es "siempre la misma", la cosa tiene poco interés... **Actualizar una predicción en función de nueva información** es mucho más relevante ("modelos" de predicción, predicción "condicional", en otros materiales).