eman ta zabal zazu

Universidad del País Vasco — Euskal Herriko Unibertsitatea

TAREA 8

EJERCICIOS

1. Queremos contrastar $H_0 : X \sim b(\theta_0)$ frente a $H_a : X \sim b(\theta \neq \theta_0)$, para lo que contamos con una m.a.s. $X_1, \ldots, X_n$.

   a) Encuentra la forma de la región crítica para el contraste localmente más potente.

   b) Haz lo propio para un contraste de Wald.

2. Comprueba que el contraste razón generalizada de verosimilitudes para la hipótesis $H_0 : X \sim N(\mu, \sigma_0^2)$ frente a $H_a : X \sim N(\mu_a \neq \mu_0, \sigma_0^2)$ (suponemos σ_0^2 conocida) es el habitualmente empleado, y hace uso del estadístico de contraste $(\overline{X} - \mu_0)/\sigma_0$ (o alguna transformación equivalente del mismo).

3. Obtén el contraste localmente más potente (o "score test") para la hipótesis $H_0 : \vec{\mu} = \vec{\mu}_0$ frente a la alternativa $H_a : \vec{\mu} \neq \vec{\mu}_0$, en el caso de una distribución normal con matriz de covarianzas Σ conocida. Comprueba que, en este caso, la distribución asintótica es también aplicable de modo exacto en pequeñas muestras.

4. *(un problema simple de aleatorización)* Genera sendos vectores con $n_1 = n_2 = 50$ observaciones, misma varianza unitaria y medias respectivas $\mu_1 = 5{,}0$ y $\mu_2 = 5{,}3$ (o cualesquiera otros valores que quieras).

   Contrastar la hipótesis $H_0 : \mu_1 = \mu_2$ es un problema muy simple, que puedes resolver con ayuda de lo que aprendiste en Estadística II. Hazlo, para comparar con los resultados que obtendrás a continuación, como si no conocieras la varianza (es decir, haz un contraste t de Student). A continuación, resuelve el mismo problema haciendo un contraste de aleatorización así:

   a) Calcula el valor t_{obs} de un estadístico conveniente —por ejemplo, $T = |\overline{X}_1 - \overline{X}_2|$— con ayuda de las dos muestras proporcionadas.

   b) Genera N permutaciones (N puede ser, por ejemplo, 200 (ó 300, ó 500, ó 1000; lo más que puedas sin perjudicar a otros usuarios de la máquina) de los $n_1 + n_2$ valores muestrales. Toma los n_1 primeros para computar $\overline{x}_1^{(1)}$ y los n_2 siguientes para computar $\overline{x}_2^{(1)}$, Computa ahora $t^{(1)} = |\overline{x}_1^{(1)} - \overline{x}_2^{(1)}|$.

   c) Repite el paso anterior para obtener $t^{(2)}, \ldots, t^{(N)}$ seudo-realizaciones adicionales de T.

   d) $t^{(1)}, \ldots, t^{(N)}$ proporcionan ahora una idea de los valores de T que sería probable obtener si realmente las $n_1 + n_2$ observaciones procedieran de la misma población. Si t_{obs} cayera entre el $100\alpha\,\%$ de mayores valores seudo-muestrales, tendrías motivos para rechazar la hipótesis al nivel de significación α.

   Nota que no has hecho ninguna hipótesis sobre la distribución de las observaciones; esencialmente el mismo procedimiento te permitiría construir contrastes en casos en que la distribución del estadístico de contraste que quieres emplear es difícil o imposible de obtener.

Observaciones, ayudas, comentarios.

1. No es posible, y de serlo no sería práctico, calcular todos los posibles $\binom{100}{50}$ valores del estadístico T (algunos acaso repetidos). La generación de unos cientos o miles de ellos es habitualmente suficiente.

2. Te serán de utilidad las funciones de S-PLUS que toman muestras al azar (como `sample()`), ordenan (`sort()`) o generan órdenaciones (como `order()`). Hay múltiples modos de hacer lo mismo; emplea el que creas mejor.

Lectura recomendada. Todos los manuales que se citan a continuación son de interés: [1], [2], [4], [6]. El libro [8] es una referencia ya clásica, pero de nivel superior. Hay problemas resueltos en todos ellos y en [5], [10] y [3].

Puedes encontrar información específica sobre contrastes de aleatorización y, en general, intensivos en computación en [9] (muy sencillo) y [7].

Referencias

[1] P. J. Bickel and K. A. Doksum. *Mathematical Statistics*. Holden-Day, Inc., San Francisco, 1977.

[2] D. R. Cox and D. V. Hinkley. *Theoretical Statistics*. Chapman and Hall, London, 1979 edition, 1974.

[3] D. R. Cox and D. V. Hinkley. *Problems and Solutions in Theoretical Statistics*. Chapman and Hall, London, 1980 edition, 1980.

[4] E.J. Dudewicz and S.N. Mishra. *Modern Mathematical Statistics*. Wiley, 1988.

[5] A. Garín and F. Tusell. *Problemas de Probabilidad e Inferencia Estadística*. Ed. Tébar-Flores, Madrid, 1991.

[6] P.H. Garthwaite, I.T. Jolliffe, and B. Jones. *Statistical Inference*. Prentice Hall, London, 1995.

[7] P. Good. *Permutation Tests*. Springer-Verlag, 1993.

[8] E. L. Lehmann. *Testing Statistical Hypothesis*. Chapman & Hall, 2 edition, 1986.

[9] E.W. Noreen. *Computer intensive methods for testing hypotheses. An introduction.* John Wiley and Sons, 1989.

[10] J. P. Romano and A. F. Siegel. *Counterexamples in Probability and Statistics*. Wadsworth and Brooks/Cole, Monterrey, California, 1986.