

برآورد

فصل هفتم

—

استنباط آماری: نتیجه‌گیری در مورد جامعه بر اساس اطلاعات حاصل از یک نمونه‌ی تصادفی از آن جامعه
پارامتر جامعه: مشخصه عددی مجهول از متغیر مورد بررسی در جامعه که برای آن استنباط آماری انجام می‌شود

پارامترهای رایج

- میانگین جامعه
- واریانس جامعه
- نسبت جامعه

استنباط‌های آماری

- برآورد
- آزمون فرض
- بازه اطمینان

برای پارامتر مجهول θ از آماره‌ی مناسبی مانند $T(X_1, \dots, X_n)$ استفاده می‌شود

اریبی برآوردگر: اختلاف میانگین برآوردگر و پارامتر $\text{bias}(T) = E[T(X_1, \dots, X_n)] - \theta$

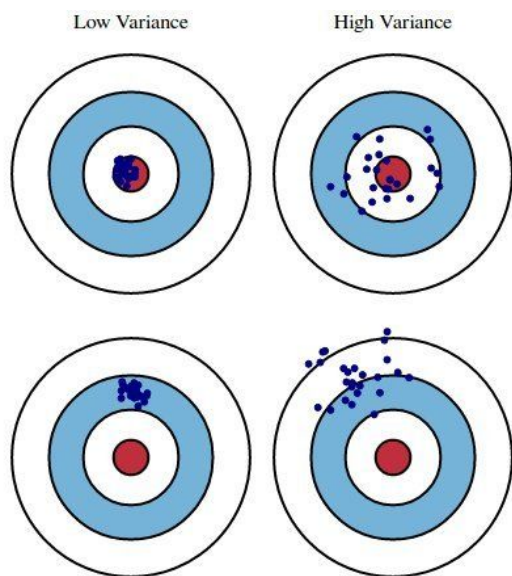
- اریبی مثبت یا بیش‌برآوردی: برآوردگر به‌طور متوسط مقداری بیشتر از مقدار پارامتر را اختیار می‌کند
- اریبی منفی یا کم‌برآوردی: برآوردگر به‌طور متوسط مقداری کمتر از مقدار پارامتر را اختیار می‌کند
- ناریبی: برآوردگر به‌طور متوسط مقداری برابر با مقدار پارامتر را اختیار می‌کند

ناریبی یک ویژگی مطلوب برای برآوردگر است

کوچک بودن واریانس هم یک ویژگی مطلوب برای برآوردگر است

واریانس کمتر: پراکندگی کمتر آن حول میانگین و در نتیجه دقت بیشتر است

برآوردگر سازگار: با افزایش حجم نمونه، واریانس به صفر میل می‌کند



برآورد میانگین و واریانس جامعه

عبداله جلیلیان، گروه آمار، دانشگاه رازی

$$\hat{\mu} = \bar{X}$$

برآورد میانگین جامعه با میانگین نمونه

- میانگین نمونه یک برآوردگر نااریب برای میانگین جامعه است
- در بین همه ی برآوردگرهای نااریب برای میانگین جامعه نرمال، میانگین نمونه کمترین واریانس را دارد
- میانگین نمونه برآوردگری سازگار برای میانگین جامعه است

$$\hat{\sigma}^2 = S^2$$

برآورد واریانس جامعه با واریانس نمونه

- واریانس نمونه یک برآوردگر نااریب برای واریانس جامعه است
- واریانس نمونه برآوردگری سازگار برای واریانس جامعه نرمال است

آزمون فرض آماری

فصل هشتم

آزمون فرض آماری

عبداله جلیلیان، گروه آمار، دانشگاه رازی

فرض آماری: گزاره‌ای در مورد توزیع جامعه‌ی آماری است

مثال:

- میانگین وزن بلوک‌های نیوجرسی یک تولیدکننده برابر با ۱۸۱۴ کیلوگرم است $\mu = 1814$
- میانگین طول عمر لامپ‌های تولید یک کارخانه دست کم ۱۲۰۰۰ ساعت است $\mu \geq 12000$
- نسبت جوش کارهای ناقص در یک سازه‌ی فلزی حداکثر ۵ درصد است $p \leq 0.05$
- واریانس مقاومت کششی میلگردهای تولید کننده الف از واریانس مقاومت کششی میلگردهای تولید کننده ب کمتر است $\sigma_1^2 < \sigma_2^2$

فرض صفر: فرض خنثی‌ای در مورد جامعه که اغلب ادعای خاص یا جدیدی را مطرح نمی‌کند

فرض مقابل: فرضی که اغلب ادعای خاص یا جدیدی را در مورد جامعه مطرح می‌کند

$$\begin{cases} H_0 : \text{Null hypothesis} \\ H_1 : \text{Alternative hypothesis} \end{cases}$$

آزمون فرض آماری

عبداله جلیلیان، گروه آمار، دانشگاه رازی

آزمون فرض آماری: اقدام به رد یا عدم رد فرض صفر در مقابل فرض مقابل بر اساس مشاهدات نمونه

$$\delta(X_1, \dots, X_n) = \begin{cases} \text{reject } H_0 \\ \text{do not reject } H_0 \end{cases}$$

آماره‌ی آزمون: آماره‌ای که بر اساس مقادیر آن تصمیم به رد یا عدم رد فرض صفر گرفته می‌شود

$$\begin{array}{ll} \text{if } T(X_1, \dots, X_n) \in C & \text{then reject } H_0 \\ \text{if } T(X_1, \dots, X_n) \notin C & \text{then do not reject } H_0 \end{array}$$

ناحیه‌ی بحرانی: ناحیه‌ای که قرار گرفتن آماره‌ی آزمون در آن منجر به رد فرض صفر می‌شود

ناحیه‌ی پذیرش: ناحیه‌ای که قرار گرفتن آماره‌ی آزمون در آن منجر به رد فرض صفر نمی‌شود

خطای نوع اول و دوم

در مسئله‌ی آزمون فرض آماری ممکن است تصمیم درست اتخاذ شود یا مرتکب دو نوع خطا شویم

فرض صفر واقعاً درست است	فرض صفر واقعاً درست نیست	
فرض صفر بر اساس مشاهده‌های نمونه رد می‌شود	خطای نوع اول: رد به غلط فرض صفر	تصمیم درست: رد صحیح فرض صفر
فرض صفر بر اساس مشاهده‌های نمونه رد نمی‌شود	تصمیم درست: عدم رد صحیح فرض صفر	خطای نوع دوم: عدم رد به غلط فرض صفر

با حجم نمونه‌ی ثابت، کاهش احتمال خطای نوع اول به افزایش احتمال خطای نوع دوم منجر می‌شود و به عکس

احتمال خطای نوع اول با یک مقدار از پیش تعیین شده توسط پژوهشگر ثابت در نظر گرفته می‌شود

مقدارهای رایج برای احتمال خطای نوع اول (سطح آزمون): ۵ درصد، ۱ درصد و ۱۰ درصد