Stat 451
3-13-18

# Prediction Intervals

①

Suppose we have sampled $n$ items from a normal population with mean $\mu$ & std. dev $\sigma$.

Goal: predict the value of a new observation & give a 95% prediction interval

Let $X$ be the new observation.

Consider the random variable $X - \bar{X}$

This will still have a normal distribution.

②

$$E[X-\bar{x}] = E[X] - E[\bar{x}]$$
$$= \mu - \mu = 0$$

$$V[X-\bar{x}] = V[x] + V[\bar{x}]$$
$$= \sigma^2 + \frac{\sigma^2}{n} = \sigma^2\left(1+\frac{1}{n}\right)$$

So $X-\bar{x} \sim N\left(0, \sigma^2\left(1+\frac{1}{n}\right)\right)$

$$\frac{X-\bar{x}-0}{\sqrt{\sigma^2\left(1+\frac{1}{n}\right)}} \sim N(0,1)$$

(3)

$$\underbrace{P\left(-1.96 < \frac{X-\bar{X}}{\sqrt{\sigma^2(1+\frac{1}{n})}} < 1.96\right)}= .95$$

$$\bar{x} - 1.96\,\sigma\sqrt{1+\tfrac{1}{n}} < X < \bar{x} + 1.96\,\sigma\sqrt{1+\tfrac{1}{n}}$$

That is, $\bar{X} \pm Z_{\alpha/2}\,\sigma\sqrt{1+\frac{1}{n}}$ is a <u>prediction</u> <u>interval</u> for $X$

Usually, use $\bar{X} \pm t_{\alpha/2}\, s\sqrt{1+\frac{1}{n}}$ df$=n-1$

(4)

Exam topics:
- Central Limit Theorem
- probabilities concerning $\bar{X}$
- Confidence intervals + sample size
- MoM & MLE
- prediction interval

| | | Assumptions |
|---|---|---|
| C.I. for $\mu$: | $\bar{x} \pm t\frac{s}{\sqrt{n}}$ df$=n-1$ | The population is normal OR $n$ is large |
| C.I. for $p$: | $\hat{p} \pm z\sqrt{\frac{\hat{p}\hat{q}}{n}}$ | — |
| C.I. for | $\sigma^2$: $\frac{(n-1)s^2}{\chi^2_{\alpha/2}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}}}$<br>$\sigma$: $\sqrt{\ldots} < \sigma < \sqrt{\ldots}$ | normal population |

(5)

| | |
|---|---|
| C.I. for $\mu_1-\mu_2$: $\bar{x}_1-\bar{x}_2 \pm t\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$, df = long formula | Each of the 2 populations must either be normal or its $n$ must be large |
| C.I. for $p_1-p_2$: $\hat{p}_1-\hat{p}_2 \pm z\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1}+\frac{\hat{p}_2\hat{q}_2}{n_2}}$ | — |
| C.I. for $\mu_d$: $\bar{d} \pm t\frac{s_d}{\sqrt{n}}$, $df = n-1$ | The population of differences is normal OR $n$ is large |
| P.I. for $X$: $\bar{x} \pm t s\sqrt{1+\frac{1}{n}}$, $df = n-1$ | normal population |

(6)

Example from HW

$$f(x) = \frac{x}{\theta^2} e^{-\frac{x^2}{2\theta^2}} \qquad x > 0$$

You were given that $\mu = \theta\sqrt{\frac{\pi}{2}}$

MOM: Set $\bar{x} = \mu$ & solve for $\theta$

$$\bar{x} = \theta\sqrt{\frac{\pi}{2}} \quad \therefore \quad \hat{\theta} = \sqrt{\frac{2}{\pi}}\,\bar{x}$$

MLE: $\mathcal{L}(\theta) = \prod_{i=1}^{n} f(x_i) = \prod_{i=1}^{n} \frac{x_i}{\theta^2} e^{-\frac{x_i^2}{2\theta^2}}$

$$= \frac{1}{\theta^{2n}} \prod_{i=1}^{n} x_i \, e^{-\frac{1}{2\theta^2}\sum x_i^2}$$ 7

$$\ell(\theta) = \ln \mathcal{L}(\theta)$$
$$= -2n \ln\theta + \sum_{i=1}^{n} \ln x_i - \frac{1}{2\theta^2}\sum x_i^2$$
$$\ell'(\theta) = -\frac{2n}{\theta} - \frac{1}{2}\sum x_i^2 (-2)\theta^{-3} \overset{\text{set}}{=} 0$$

$$-2n\theta^2 + \sum x_i^2 = 0$$
$$\theta^2 = \frac{\sum x_i^2}{2n} \qquad \therefore \hat{\theta} = \sqrt{\frac{\sum x_i^2}{2n}}$$

8

For MoM, what if $\mu$ had not been given?

$$\mu = E[X] = \int x f(x)\,dx$$
$$= \int_0^\infty x \frac{x}{\theta^2} e^{-\frac{x^2}{2\theta^2}}\,dx$$
$$= \int_0^\infty \frac{x^2}{\theta^2} e^{-\frac{x^2}{2\theta^2}}\,dx$$
$$= \cancel{-xe^{-x^2/2\theta^2}\Big|_0^\infty} + \int_0^\infty e^{-\frac{x^2}{2\theta^2}}\,dx$$

$u = x$

$dv = \frac{x}{\theta^2} e^{-\frac{x^2}{2\theta^2}}\,dx$

$du = dx$

$v = -e^{-x^2/2\theta^2}$

$$= \int_0^\infty e^{-\frac{1}{2}z^2} \theta \, dz \qquad \text{let } z = \frac{x}{\theta}, \quad dz = \frac{1}{\theta} dx$$

$$= \theta \sqrt{2\pi} \underbrace{\int_0^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz}_{.5} = \theta\sqrt{2\pi} \cdot \frac{1}{2} = \theta\sqrt{\frac{\pi}{2}}$$

---

Exam Thurs 3/15

1 page of notes, front & back

$z$, $t$, $\chi^2$ tables

calculator or laptop