# Utledning av prediksjonsintervall

- Situasjon: $X_1, X_2, \ldots, X_n$ tilfeldig utvalg fra $n(x; \mu, \sigma)$-populasjonen
  - Ønsker prediksjonsintervall for ny observasjon $X_0 \sim n(x_0; \mu, \sigma)$

- Har at

$$\bar{X} \sim n\left(\bar{x}; \mu, \sqrt{\frac{\sigma^2}{n}}\right)$$

- Siden $X_0$ er uavhengig av $\bar{X}$ har vi også

$$X_0 - \bar{X} \sim n\left(z; 0, \sqrt{\sigma^2 + \frac{\sigma^2}{n}}\right)$$

## Utledning av prediksjonsintervall

- Situasjon: $X_1, X_2, \ldots, X_n$ tilfeldig utvalg fra $n(x; \mu, \sigma)$-populasjonen
  - Ønsker prediksjonsintervall for ny observasjon $X_0 \sim n(x_0; \mu, \sigma)$
- Har at

$$\bar{X} \sim n\left(\bar{x}; \mu, \sqrt{\frac{\sigma^2}{n}}\right)$$

- Siden $X_0$ er uavhengig av $\bar{X}$ har vi også

$$X_0 - \bar{X} \sim n\left(z; 0, \sqrt{\sigma^2 + \frac{\sigma^2}{n}}\right)$$

- Ved å standardisere får vi

$$\frac{X_0 - \bar{X}}{\sqrt{\sigma^2\left(1 + \frac{1}{n}\right)}} \sim n(z; 0, 1)$$

# Utledning av prediksjonsintervall

- ⋆ Situasjon: $X_1, X_2, \ldots, X_n$ tilfeldig utvalg fra $n(x; \mu, \sigma)$-populasjonen
  - Ønsker prediksjonsintervall for ny observasjon $X_0 \sim n(x_0; \mu, \sigma)$

- ⋆ Har at

$$\bar{X} \sim n\left(\bar{x}; \mu, \sqrt{\frac{\sigma^2}{n}}\right)$$

- ⋆ Siden $X_0$ er uavhengig av $\bar{X}$ har vi også

$$X_0 - \bar{X} \sim n\left(z; 0, \sqrt{\sigma^2 + \frac{\sigma^2}{n}}\right)$$

- ⋆ Ved å standardisere får vi

$$\frac{X_0 - \bar{X}}{\sqrt{\sigma^2\left(1 + \frac{1}{n}\right)}} \sim n(z; 0, 1)$$

- ⋆ Nå to muligheter: Verdien til $\sigma^2$ er kjent eller ukjent.

# $\chi^2$ og $t$-fordeling: Noen resultater

- ⋆ Når $X_1, X_2, \ldots, X_n$ er tilfeldig utvalg fra $n(x; \mu, \sigma)$ er

$$\frac{(n-1)S^2}{\sigma^2} = \sum_{i=1}^{n} \left( \frac{X_i - \bar{X}}{\sigma} \right)^2 \sim \chi^2_{n-1}$$

- ⋆ Teorem: La $Z \sim n(z; 0, 1)$ og $V \sim \chi^2_\nu$ være uavhengige. Da gjelder

$$T = \frac{Z}{\sqrt{\frac{V}{\nu}}} \sim t_\nu$$

# Utledning av konfidensintervall

- ⋆ Situasjon: Anta $X_1, X_2, \ldots, X_n$ tilfeldig utvalg fra $f(x;\theta)$-populasjonen. Ønsker $(1-\alpha)\cdot 100\%$ konfidensintervall for $\theta$.

1) Estimator for $\theta$: $\widehat{\theta}$

2) Sett $Z = h(\widehat{\theta}, \theta)$ der $h(\cdot,\cdot)$ er en funksjon (uten andre ukjente parametre enn $\theta$) slik at $Z$ har en kjent sannsynlighetsfordeling (uten ukjente parametre).

3) Dermed har vi at

$$P\left(z_{1-\frac{\alpha}{2}} \leq h(\widehat{\theta}, \theta) \leq z_{\frac{\alpha}{2}}\right) = 1-\alpha.$$

4) Løs hver ulikhet med hensyn på $\theta$ (hver for seg), og sett deretter de to ulikhetene sammen igjen med $\theta$ i midten slik at man får

$$P\left(\widehat{\theta}_L(X_1, X_2, \ldots, X_n) \leq \theta \leq \widehat{\theta}_U(X_1, X_2, \ldots, X_n)\right) = 1-\alpha.$$

5) Konkluder: Et $(1-\alpha)\cdot 100\%$ konfidensintervall for $\theta$ er

$$\left[\widehat{\theta}_L(X_1, X_2, \ldots, X_n), \widehat{\theta}_U(X_1, X_2, \ldots, X_n)\right].$$

6) Sett inn observerte tall for $X_1, X_2, \ldots, X_n$ og få et numerisk intervall.