NTNU
Kunnskap for ei betre verd

Konfidensintervall

Torstein Fjeldstad
Institutt for matematiske fag, NTNU
14.03.2019

I dag

- Repetisjon
- Konfidensintervall for p i binomisk fordeling
- Konfidensintervall for σ^2 i normalfordelinga
- Prediksjonsintervall

Repetisjon I

Kji-kvadratfordeling (χ^2-fordeling)

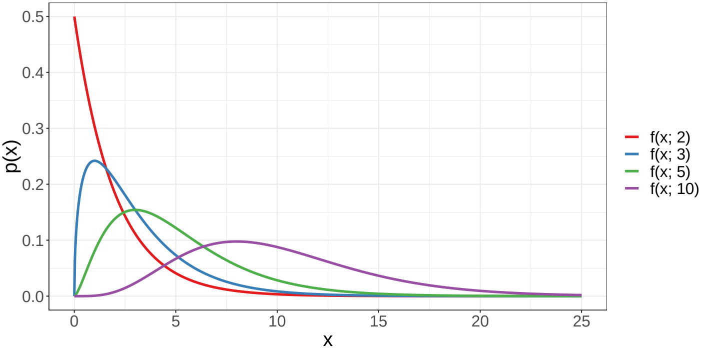
Sannsynstettleiken til ein χ^2-fordelt variabel X er:

$$f(x;\nu) = \frac{1}{2^{\nu/2}\Gamma(\frac{\nu}{2})} x^{\frac{\nu}{2}-1} e^{-\frac{x}{2}}.$$

Det kan visast at

$$E(X) = \nu \qquad \text{og} \qquad Var(X) = 2\nu.$$

Repetisjon II



Repetisjon III

Utvalsfordelinga til S^2

Dersom $X_1, X_2, \ldots, X_n$ er uavhengige og $X_i \sim n(x_i; \mu, \sigma)$ er

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

og $\frac{(n-1)S^2}{\sigma^2}$ og $\bar{X}$ er uavhengige.

Repetisjon IV

Student t-fordelinga

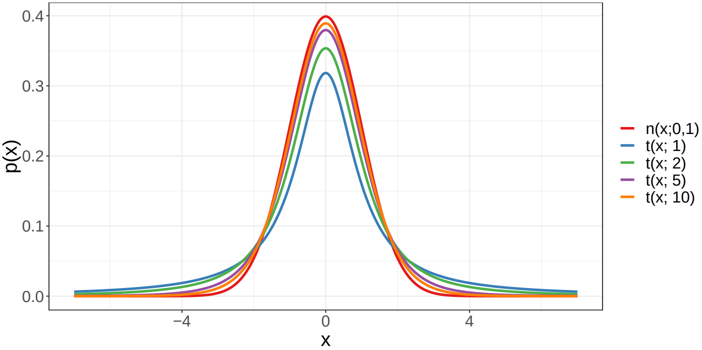
La $Z \sim n(z; 0, 1)$ og $V \sim \chi^2_\nu$ der Z og V er uavhengige. Då vil fordelinga

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

ha sannsynstettleik

$$f(t; \nu) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sqrt{\pi\nu}} \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}} \qquad -\infty < t < \infty.$$

Repetisjon V



Utledning av konfidensintervall

Situasjon: Anta $X_1, X_2, \ldots, X_n$ tilfeldig utval frå $f(x;\theta)$-populasjonen. Ynskjer eit $(1-\alpha)\cdot 100\%$ konfidensintervall for θ.

1. Finn ein estimator for θ, $\hat{\theta}$ (t.d. SME)
2. La $Z = h(\hat{\theta}, \theta)$ der $h(\cdot, \cdot)$ er ein funksjon s.a. Z har ei kjend fordeling.
3. Då har me at

$$P(z_{1-\alpha/2} \leq h(\hat{\theta}, \theta) \leq z_{\alpha/2}) = 1 - \alpha$$

4. Løys ulikskapane (mhp. θ) kvar for seg og finn eit uttrykk med θ i midten

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

5. Et $(1-\alpha)\cdot 100$ % konfidensintervall for θ er

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

Kva situasjonar har me sett så langt?

1. $X_1, X_2, \ldots, X_n$ uif $n(x; \mu, \sigma)$ (σ kjend)

$$\left[\bar{X} - z_{\alpha/2}\sqrt{\frac{\sigma^2}{n}}, \bar{X} + z_{\alpha/2}\sqrt{\frac{\sigma^2}{n}}\right]$$

2. $X_1, X_2, \ldots, X_n$ uif $n(x; \mu, \sigma)$ (σ ukjend)

$$\left[\bar{X} - t_{\alpha/2,n-1}\sqrt{\frac{S^2}{n}}, \bar{X} + t_{\alpha/2,n-1}\sqrt{\frac{S^2}{n}}\right]$$

Konfidensintervall for p i binomisk fordeling

Eksempel

Anta at av eit vareparti på 100 TV-skjermar har du observert at 5 av desse er defekte. Anta at sannsynet for at ein tilfeldig valgt TV ikkje verkar er konstant og TV-ane verkar uavhengig av kvarandre. Finn eit $(1-\alpha)\cdot 100$ % konfidensintervall for sannsynet for at ein tilfeldig valgt TV verkar.

Konfidensintervall for p i binomisk fordeling

Eksempel

Anta at av eit vareparti på 100 TV-skjermar har du observert at 5 av desse er defekte. Anta at sannsynet for at ein tilfeldig valgt TV ikkje verkar er konstant og TV-ane verkar uavhengig av kvarandre. Finn eit $(1-\alpha)\cdot 100$ % konfidensintervall for sannsynet for at ein tilfeldig valgt TV verkar.

1. $X_1, X_2, \ldots, X_{100}$ er bernoullifordelt med sannsyn p

Konfidensintervall for p i binomisk fordeling

Eksempel

Anta at av eit vareparti på 100 TV-skjermar har du observert at 5 av desse er defekte. Anta at sannsynet for at ein tilfeldig valgt TV ikkje verkar er konstant og TV-ane verkar uavhengig av kvarandre. Finn eit $(1-\alpha)\cdot 100$ % konfidensintervall for sannsynet for at ein tilfeldig valgt TV verkar.

1. $X_1, X_2, \ldots, X_{100}$ er bernoullifordelt med sannsyn p
2. Estimator for p: $\widehat{p} = \bar{X}$

Konfidensintervall for p i binomisk fordeling

Eksempel

Anta at av eit vareparti på 100 TV-skjermar har du observert at 5 av desse er defekte. Anta at sannsynet for at ein tilfeldig valgt TV ikkje verkar er konstant og TV-ane verkar uavhengig av kvarandre. Finn eit $(1-\alpha)\cdot 100$ % konfidensintervall for sannsynet for at ein tilfeldig valgt TV verkar.

1. $X_1, X_2, \ldots, X_{100}$ er bernoullifordelt med sannsyn p
2. Estimator for p: $\widehat{p} = \bar{X}$
3. Sentralgrenseteoremet gir

$$Z = \frac{\bar{X} - p}{\sqrt{\frac{p(1-p)}{n}}} \sim n(z; 0, 1)$$

Konfidensintervall for σ^2 i normalfordelinga

Situasjon: $X_1, X_2, \ldots, X_n$ uif $n(x; \mu, \sigma)$ der både μ og σ er ukjende.
Mål: finn eit $(1 - \alpha) \cdot 100$ % konfidensintervall for σ^2

Konfidensintervall for σ^2 i normalfordelinga

Situasjon: $X_1, X_2, \ldots, X_n$ uif $n(x; \mu, \sigma)$ der både μ og σ er ukjende.
Mål: finn eit $(1 - \alpha) \cdot 100$ % konfidensintervall for σ^2

1. Estimator for σ^2; $S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$

Konfidensintervall for σ^2 i normalfordelinga

Situasjon: $X_1, X_2, \ldots, X_n$ uif $n(x; \mu, \sigma)$ der både μ og σ er ukjende.
Mål: finn eit $(1-\alpha) \cdot 100$ % konfidensintervall for σ^2

1. Estimator for σ^2; $S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$
2. Hugs

$$V = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$$

Nye situasjonar

1. $X_1, X_2, \ldots, X_n$ uif $b(x; n, p)$ (approksimativt)

$$\left[\widehat{p} - z_{\alpha/2}\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}}, \widehat{p} + z_{\alpha/2}\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}}\right]$$

2. $X_1, X_2, \ldots, X_n$ uif $n(x; \mu, \sigma)$ (μ ukjend)

$$\left[\frac{(n-1)S^2}{\chi^2_{\alpha/2,n-1}}, \frac{(n-1)S^2}{\chi^2_{1-\alpha/2,n-1}}\right]$$

Prediksjonsintervall

Situasjon: $X_1, X_2, \ldots, X_n$ tilfeldig utval med $X_i \sim n(x_i; \mu, \sigma)$ der σ er kjend og μ ukjend. La $X_0 \sim n(x_0; \mu, \sigma)$ vere ei framtidig måling, uavhengig av $X_1, X_2, \ldots, X_n$.

— Ynskjer eit $(1-\alpha)100\%$ prediksjonsintervall for X_0

— Estimator for μ, $\hat{\mu} = \bar{X}$

— Start med

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

— Observator

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

— $(1-\alpha)100\%$ prediksjonsintervall for X_0

$$\left[\bar{X} - z_{\alpha/2}\sqrt{\sigma^2\left(\frac{1}{n} + 1\right)}, \bar{X} + z_{\alpha/2}\sqrt{\sigma^2\left(\frac{1}{n} + 1\right)}\right]$$

Skilnad konfidensintervall og prediksjonsintervall

Konfidensintervall:

$$P\left(\bar{X} - z_{\alpha/2}\sqrt{\frac{\sigma^2}{n}} \leq \mu \leq \bar{X} + z_{\alpha/2}\sqrt{\frac{\sigma^2}{n}}\right) = 1 - \alpha$$

Prediksjonsintervall:

$$P\left(\bar{X} - z_{\alpha/2}\sqrt{\sigma^2\left(\frac{1}{n} + 1\right)} \leq X_0 \leq \bar{X} + z_{\alpha/2}\sqrt{\sigma^2\left(\frac{1}{n} + 1\right)}\right) = 1 - \alpha$$

Neste veke

- To-utval
- Hypotesetesting