Plan:

1. Konfidensintervall for μ
2. Noen flere fordelinger
3. Konfidensintervall for σ^2
4. Konfidensintervall for p

Pensum:

[L] 6.3, 5.9

Repetisjon:

a) Beskrivelse av datasettet / utvalget:

Viktige begrep:
- frekvenstabell / histogram
- sentralmål: gjennomsnitt, median (i utvalg)
- kvartiler / boksplott
- spredningsmål: std avvik, kvartilbredde (i utvalg)

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

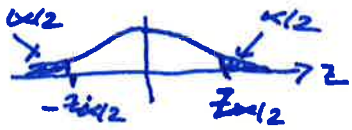
$$S_x = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2}$$

b) Konfidensintervall for μ når σ er kjent:

Anta:
- $X_1, \ldots, X_n$ er uavhengige, identisk fordelte med $E(X_i)=\mu$ og $Var(X_i)=\sigma^2$ for alle i
- $\bar{X} = \frac{X_1+\ldots+X_n}{n}$ er (tilnærmet) normalfordelt { hver X_i normalfordelt eller n stor ($n \geq 30$)
- μ skal estimeres når σ^2 er kjent

Velg: Signifikansnivå $\alpha \rightsquigarrow (1-\alpha)\cdot 100\%$ konfidensintervall

$\alpha = 0.05$ 95% konfidensintervall



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

1.96 når $\alpha = 0.05$

et std. avvik for $\bar{X}$

1 Konfidensintervall μ (σ ikke kjent)

Antar: $X_1, \dots, X_n$ uavh, id. distr. normalf. $E(x_i) = \mu$ $Var(x_i) = \sigma^2$

~~$\bar{X}$ (...) normalfordelt~~

skal estimere μ (σ^2 ukjent)

Finn: $\bar{x}, s^2$ fra utvalget

$$\boxed{\bar{x} \pm z_{\alpha/2} \cdot \sigma/\sqrt{n}}$$

$\sigma^2 \leftarrow s^2$?

Hvis n er stor: ok ($n \geq 30$)

Hvis n er liten: må bruke t-verdier istedet for z-verdier

Konfidensintervall:

$$\boxed{\bar{x} \pm t^{n-1}_{\alpha/2} \cdot s/\sqrt{n}}$$

Eks: 79 87 94 92 97

$n = 5$
$\bar{x} = 89.8$
$s^2 = 49.7 \rightarrow s = 7.0498$

$\left.\begin{array}{l}\alpha = 0.05: \\ n-1 = 4:\end{array}\right\}$ $t^{n-1}_{\alpha/2} = t^{4}_{0.025}$

Kalkulator:

4 $\boxed{INV}$ $\boxed{df, tzP}$ 0.975 $\boxed{=}$

$t^{n-1}_{\alpha/2} = 2.7764$

($z_{\alpha/2} = 1.96$)

Intervall: $\bar{x} \pm t^{n-1}_{\alpha/2} \cdot s/\sqrt{n}$

$$= 89.8 \pm 2.7764 \cdot \frac{7.0498}{\sqrt{5}}$$

$$= 89.8 \pm 8.75$$

$$= [81.05, 98.55]$$

Z-intervall m/ $\sigma = 7.05$: $89.8 \pm 1.96 \cdot \frac{7.05}{\sqrt{5}} = [83.62, 95.98]$

② Noen flere fordelinger: [L] 5.9

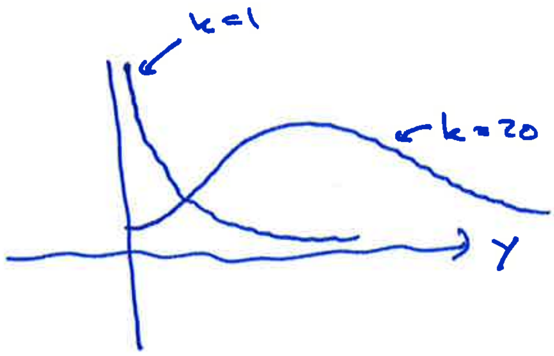
a) χ^2_k – fordelingen (χ = chi / khi)

Anta at $Z_1, Z_2, \ldots, Z_k$ er k std. normalfordelte variabler som er uavhengige, så er

$$\chi^2_k = Z_1^2 + Z_2^2 + \ldots + Z_k^2$$

χ^2_k-fordelt, og k kalles antall frihetsgrader.

$Y = Z_1^2 + Z_2^2 + \ldots + Z_k^2$:



$Y \geq 0$

$E(Y) = k \quad Var(Y) = 2k$

$f(y) = K\, y^{k/2-1} e^{-y/2}, \quad y \geq 0$

Resultat: $X_1, \ldots, X_n$ uavhengige, normalfordelte variabler med $E(X_i) = \mu$, $Var(X_i) = \sigma^2$

$\Downarrow$

$$Y = \frac{(n-1)S^2}{\sigma^2}$$ er χ^2_{n-1} - fordelt, dvs χ^2-fordelt med $n-1$ frihetsgrader

Husk: $$S^2 = \frac{1}{n-1} \cdot \sum_{i=1}^{n} (X_i - \bar{X})^2$$

$\Downarrow$

$$Y = \frac{1}{\sigma^2} \sum_{i=1}^{n} (X_i - \bar{X})^2 = \sum_{i=1}^{n} \left(\frac{X_i - \bar{X}}{\sigma} \right)^2 \longleftarrow \chi^2_{n-1}\text{- fordelt}$$

χ^2-fordeling med n frihetsgrader $\longrightarrow$ $$\sum_{i=1}^{n} Z_i^2 = \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma} \right)^2$$

b) **Student t-fordeling = t-fordeling**

Anta: Z : $N(0,1)$, dvs std. normalfordelt

U : χ^2_n - fordelt, dvs χ^2-fordelt med n frihetsgrader

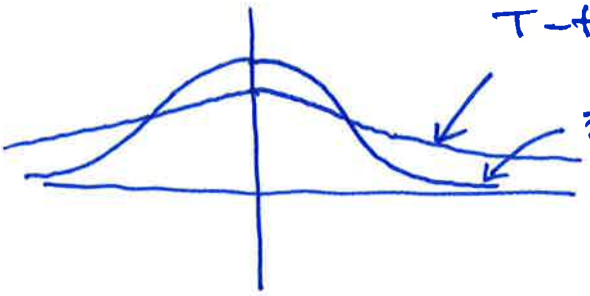
Z, U uavhengige

$\Downarrow$

Defn: $$T = \frac{Z}{\sqrt{U/n}}$$ t-fordelt med n frihetsgrader

Generelt: - nesten normalfordelt når n er stor

- tetthetsfn. $f(t) = K \cdot \left(1 + \frac{t^2}{n}\right)^{-(n+1)/2}$



$E(T) = 0$

$Var(T) > 1$

Viktig spesialtilfelle:

$X_1, \ldots, X_n$: uavhengige, normalfordelte m/ $E(X_i) = \mu$, $Var(X_i) = \sigma^2$

$\Downarrow$

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$ er t-fordelt med $(n-1)$ frihetsgrader

$$\boxed{Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1) \text{ std-normal fordelt}}$$

$\parallel$

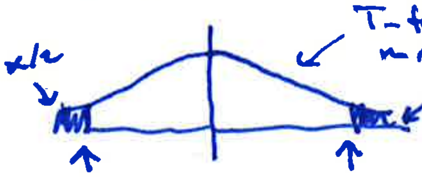
$$\frac{\dfrac{\bar{X} - \mu}{\sigma/\sqrt{n}}}{S/\sigma} = \frac{Z}{\sqrt{\underbrace{\frac{(n-1)S^2}{\sigma^2}} : (n-1)}}$$

$Y : \chi^2_{n-1}$ - fordelt

Oppsummering: Konfidensintervall for μ når σ er ukjent

Anta: $X_1, X_2, \ldots, X_n$ uavhengige, normalfordelte med $E(X_i) = \mu$ og $Var(X_i) = \sigma^2$

Velg: signifikansnivå α



$\alpha = 0{,}05$

95% konf. intervall

$t_{\alpha/2}^{n-1} = 2{,}7764$

$$\boxed{\text{Konfidensintervall: } \bar{X} \pm t_{\alpha/2}^{n-1} \cdot S/\sqrt{n}}$$

$A = \bar{X} - t_{\alpha/2}^{n-1} \cdot S/\sqrt{n}$

$B = \bar{X} + t_{\alpha/2}^{n-1} \cdot S/\sqrt{n}$

Altså: Hvis $X_1, \ldots, X_n$ er uavhengige og normalfordelte med $E(X_i) = \mu$, $Var(X_i) = \sigma^2$, så er

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

et $(1-\alpha)\cdot 100\%$ konfidensintervall for μ.

Metode konstruert for tilfeller der n er liten.

③ Konfidensintervall for σ^2:

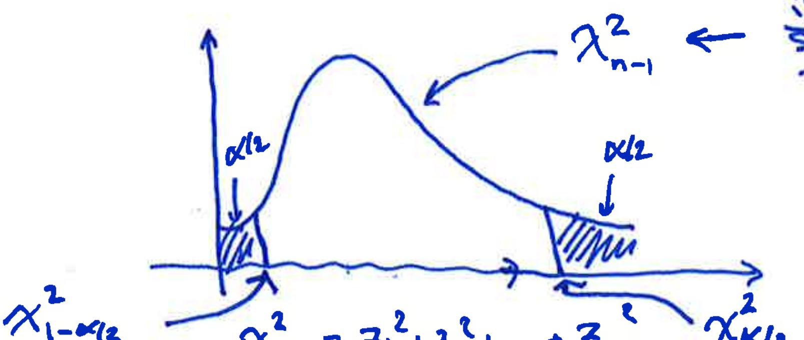
Anta: $X_1, \ldots, X_n$ uavhengige, normalfordelte variabler med $E(X_i) = \mu$, $Var(X_i) = \sigma^2$

Velg: signifikansnivå α $\longrightarrow$ $(1-\alpha)\cdot 100\%$ konfidensintervall

$\alpha = 0{,}05$ — 95% — " —

Resultat: Konfidensintervall for σ^2: $\left[\dfrac{S^2\cdot(n-1)}{\chi_{\alpha/2}},\ \dfrac{S^2(n-1)}{\chi_{1-\alpha/2}}\right]$

$$\boxed{\text{Husk: } Y = \frac{(n-1)S^2}{\sigma^2} \text{ er } \chi^2_{n-1}\text{-fordelt}}$$



$$\chi^2_{n-1} = Z_1^2 + Z_2^2 + \ldots + Z_{n-1}^2$$

$\alpha = 0{,}05$: $\qquad \chi^2_{1-\alpha/2} = 2.70$

$n = 10$: $\qquad \chi^2_{\alpha/2} = 19.02$

$df = 9 \quad (n-1)$

$\alpha/2 = 0{,}025$

Eks: 91 103 99 108 96 104 98 98 83 107

$n=10$ uavhengige, normalfordelte variabler

$n-1=9$ frihetsgrader

$\alpha=0.05$: $\chi^2_{1-\alpha/2} = 2{,}70$

$\chi^2_{\alpha/2} = 19{,}02$

$s^2 = 57.344$

Konfidensintervall for σ^2:

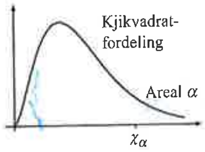
$$\left[\frac{s^2(n-1)}{\chi_{\alpha/2}} , \frac{s^2(n-1)}{\chi_{1-\alpha/2}} \right]$$

"

σ^2 : $[27.1 , 191.1]$

E.6 Kjikvadratfordelingens kvantiltabell

Tabellen viser den kritiske verdien χ_α for forskjellige valg av nivået α.



Antall frihets-grader	Areal *alfa*						Areal *alfa*					
	0,998	0,995	0,990	0,975	0,950	0,900	0,100	0,050	0,025	0,010	0,005	0,002
1	0,00	0,00	0,00	0,00	0,00	0,02	2,71	3,84	5,02	6,63	7,88	9,55
2	0,00	0,01	0,02	0,05	0,10	0,21	4,61	5,99	7,38	9,21	10,60	12,43
3	0,04	0,07	0,11	0,22	0,35	0,58	6,25	7,81	9,35	11,34	12,84	14,80
4	0,13	0,21	0,30	0,48	0,71	1,06	7,78	9,49	11,14	13,28	14,86	16,92
5	0,28	0,41	0,55	0,83	1,15	1,61	9,24	11,07	12,83	15,09	16,75	18,91
6	0,49	0,68	0,87	1,24	1,64	2,20	10,64	12,59	14,45	16,81	18,55	20,79
7	0,74	0,99	1,24	1,69	2,17	2,83	12,02	14,07	16,01	18,48	20,28	22,60
8	1,04	1,34	1,65	2,18	2,73	3,49	13,36	15,51	17,53	20,09	21,95	24,35
9	1,37	1,73	2,09	2,70	3,33	4,17	14,68	16,92	19,02	21,67	23,59	26,06
10	1,73	2,16	2,56	3,25	3,94	4,87	15,99	18,31	20,48	23,21	25,19	27,72
11	2,13	2,60	3,05	3,82	4,57	5,58	17,28	19,68	21,92	24,73	26,76	29,35
12	2,54	3,07	3,57	4,40	5,23	6,30	18,55	21,03	23,34	26,22	28,30	30,96
13	2,98	3,57	4,11	5,01	5,89	7,04	19,81	22,36	24,74	27,69	29,82	32,54
14	3,44	4,07	4,66	5,63	6,57	7,79	21,06	23,68	26,12	29,14	31,32	34,09
15	3,92	4,60	5,23	6,26	7,26	8,55	22,31	25,00	27,49	30,58	32,80	35,63
16	4,41	5,14	5,81	6,91	7,96	9,31	23,54	26,30	28,85	32,00	34,27	37,15
17	4,92	5,70	6,41	7,56	8,67	10,09	24,77	27,59	30,19	33,41	35,72	38,65
18	5,44	6,26	7,01	8,23	9,39	10,86	25,99	28,87	31,53	34,81	37,16	40,14
19	5,97	6,84	7,63	8,91	10,12	11,65	27,20	30,14	32,85	36,19	38,58	41,61
20	6,51	7,43	8,26	9,59	10,85	12,44	28,41	31,41	34,17	37,57	40,00	43,07
21	7,07	8,03	8,90	10,28	11,59	13,24	29,62	32,67	35,48	38,93	41,40	44,52
22	7,64	8,64	9,54	10,98	12,34	14,04	30,81	33,92	36,78	40,29	42,80	45,96
23	8,21	9,26	10,20	11,69	13,09	14,85	32,01	35,17	38,08	41,64	44,18	47,39
24	8,80	9,89	10,86	12,40	13,85	15,66	33,20	36,42	39,36	42,98	45,56	48,81
25	9,39	10,52	11,52	13,12	14,61	16,47	34,38	37,65	40,65	44,31	46,93	50,22
26	9,99	11,16	12,20	13,84	15,38	17,29	35,56	38,89	41,92	45,64	48,29	51,63
27	10,60	11,81	12,88	14,57	16,15	18,11	36,74	40,11	43,19	46,96	49,65	53,02
28	11,21	12,46	13,56	15,31	16,93	18,94	37,92	41,34	44,46	48,28	50,99	54,41
29	11,83	13,12	14,26	16,05	17,71	19,77	39,09	42,56	45,72	49,59	52,34	55,79
30	12,46	13,79	14,95	16,79	18,49	20,60	40,26	43,77	46,98	50,89	53,67	57,17
31	13,10	14,46	15,66	17,54	19,28	21,43	41,42	44,99	48,23	52,19	55,00	58,54
32	13,73	15,13	16,36	18,29	20,07	22,27	42,58	46,19	49,48	53,49	56,33	59,90
33	14,38	15,82	17,07	19,05	20,87	23,11	43,75	47,40	50,73	54,78	57,65	61,26
34	15,03	16,50	17,79	19,81	21,66	23,95	44,90	48,60	51,97	56,06	58,96	62,61
35	15,69	17,19	18,51	20,57	22,47	24,80	46,06	49,80	53,20	57,34	60,27	63,95
40	19,03	20,71	22,16	24,43	26,51	29,05	51,81	55,76	59,34	63,69	66,77	70,62
45	22,48	24,31	25,90	28,37	30,61	33,35	57,51	61,66	65,41	69,96	73,17	77,18
50	26,01	27,99	29,71	32,36	34,76	37,69	63,17	67,50	71,42	76,15	79,49	83,66
60	33,27	35,53	37,48	40,48	43,19	46,46	74,40	79,08	83,30	88,38	91,95	96,40
70	40,75	43,28	45,44	48,76	51,74	55,33	85,53	90,53	95,02	100,43	104,21	108,93
80	48,40	51,17	53,54	57,15	60,39	64,28	96,58	101,88	106,63	112,33	116,32	121,28
100	64,11	67,33	70,06	74,22	77,93	82,36	118,50	124,34	129,56	135,81	140,17	145,58

Tabellverdiene er beregnet med Excel-funksjonen INVERS.KJI.FORDELING(alfa;frihetsgrad).

For et høyere antall frihetsgrader (n) kan du benytte formelen $\chi_\alpha = n + z_\alpha\sqrt{2n}$, der z_α er den tilsvarende kritiske verdien for normalfordelingen (se tabell E.4).

④ Konfidensintervall for p

X : binomisk fordelt med parametre (n, p)
$\|$ med $\underline{n \text{ stor}}$ slik at X er tilnærmet normalfordelt
$X_1 + X_2 + \dots + X_n$ $\quad \hat{p} = \frac{X}{n}$

$$\boxed{\underline{\text{Konfidensintervall:}} \quad \left[\hat{p} - z_{\alpha/2} \cdot \sqrt{\frac{\hat{p} \cdot (1-\hat{p})}{n}} \; , \; \hat{p} + z_{\alpha/2} \cdot \sqrt{\frac{\hat{p} \cdot (1-\hat{p})}{n}} \right]}$$

$\underline{\text{Husk:}}$ $Var(\hat{p}) = Var\left(\frac{X}{n}\right) = \frac{1}{n^2} \cdot Var(X) = \frac{\cancel{n}pq}{n^{\cancel{2}}} = \frac{pq}{n}$

$$= \frac{p \cdot (1-p)}{n}$$

$Var(\bar{x}) = \sigma / \sqrt{n}$

$\underline{\text{Hvor stor må } n \text{ være?}}$ $\quad npq \geq 5$

$n \cdot \hat{p}\,(1-\hat{p}) \geq 5$ $\quad n \cdot p(1-p) \geq 5$

$\underline{\text{Eks:}}$ $n = 200$ $\quad \hat{p} = \frac{34}{200} = 0.17$
$X = 34$
$\alpha = 0.05$ $\quad 200 \cdot 0.17 \cdot 0.83 = 14.11 \geq 5$ ok

Konfidens-intervall: (95%) $\quad \hat{p} \pm z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.17 \pm 1.96 \cdot \sqrt{\frac{0.17 \cdot 0.83}{200}}$

$$= 0.17 \pm 0.052$$

$\underline{[0.118, 0.222]}$