Statistikk for ingeniører

SISTE digitale plenumstime - uke 10 (25.10.2021)

Mette Langaas, Institutt for matematiske fag, NTNU

Fellesmodul

Uke 10: Oppsummering og skoleeksamen

7-9: Statistisk inferens

Uke 7-8: Estimering og hypotesetesting

Uke 9: Lineær regresjon

2-6: Sannsynlighetsregning

Uke 2-3: Sannsynlighetsteori

Uke 4-6: Sannsynlighets-modeller

Uke 1: Beskrivende statistikk

Pensum fellesmodul

Pensum i fellesmodulen defineres utifra kapitler i læreboka: Statistikk for universiteter og høgskoler av Gunnar G. Løvaas.

Uke 1: Introduksjon og beskrivende statistikk

Kapittel 1 og kapittel 2.1 - 2.5, 7.2 og 7.3.2

Uke 2: Hendelser og sannsynlighet

Kapittel 3.1 - 3.6

Uke 3: Stokastiske variabler

Kapittel 4

Uke 4, 5 og 6: Vanlige sannsynlighetsmodeller

Kapittel 5.1 - 5.2, 5.4 - 5.8, 5.9.2, samt jupyter notatbok om weibullfordeling og systempålitelighet

Uke 7: Estimering og konfidensintervall

Kapittel 6.1 - 6.3 (utenom 6.2.6 og 6.3.4)

Uke 8: Hypotesetesting

Kapittel 6.4 og 6.5

Uke 9: Lineær regresjon

Kapittel 7.1 - 7.3

Plan for timen

Med fokus i formelarket og tabeller

↓

Forklare hvilke spørsmål/tema disse løser

↓

Se på noen enkle eksempler og vise til eksamensoppgaver

↓

Minne om ressurser på Bb

↓

Noen råd for eksamensdagen og vi leser forsiden sammen!

Kandidat ID
? står att

i **Institutt for matematiske fag**

Eksamensoppgåve i ISTA1001, ISTA1002, ISTA1003, ISTG1001, ISTG1002, ISTG1003, ISTT1001, ISTT1002, ISTT1003, VB6200 Statistikk

Eksamensdato: 29.10.2021

Eksamenstid (frå-til): 09:00 – 12:00

Hjelpemiddelkode/Tillatne hjelpemiddel: C
Godkjend kalkulator
Digitale ressursar under eksamen: Formelark og fem tabellar (binomisk kumulativfordeling, Poisson kumulativfordeling, normal kumulativfordling, normalfordeling kritisk verdi og t-fordeling kritisk verdi) er lagt ved eksamen som pdf-filer

Fagleg kontakt under eksamen:
Mette Langaas (988 47 649)
Ketil Arnesen (952 93 103)

Tabell: binomisk kumulativ s…
Tabell: Poisson kumulativ sa…
Tabell: normal kumulativ san…
Tabell: kritisk verdi i normalfo…
Tabell: kritisk verdi i t-fordelin…
Formelark

1 Beskrivende statistikk

Mål: bli kjent med begreper!

Gjennomsnitt

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

Empirisk varians

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

Empirisk standardavvik

$$s = \sqrt{s^2}$$

Empirisk korrelasjon

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

Minste kvadratsums rette linje $y = a + bx$

$$b = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

$$a = \bar{y} - b\bar{x}.$$

Eksamensdata

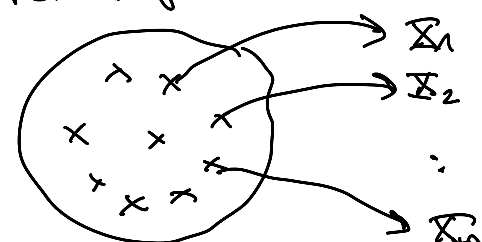
	1	2	..	7
1	15			
2	10			
:	:			
n	7			

gjsn
sd

gitt et datasett
↓
beskrive
⇓
gått videre

Populasjon

Utvalg



levetid
innhold fe

2 Hendelser og sannsynlighet

ALLTID med på eksamen i et hvert grunnemne i statistikk!

Addisjonsregelen

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Betinget sannsynlighet

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Total sannsynlighet La hendelsene A_1, A_2, ..., A_n danne en partisjon (oppdeling) av utfallsrommet. Da er

$$P(B) = \sum_{i=1}^{n} P(B|A_i)P(A_i)$$

Bayes regel

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Uavhengige hendelser

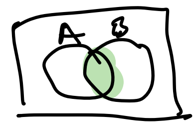
$$P(A \cap B) = P(A) \cdot P(B)$$
$$P(A|B) = P(A),\ P(B|A) = P(B)$$

Utfallsrom $S = \{M, K\}$

$S = \{1, 2, 3, 4, 5, 6\}$

Hendelser: A: minst 3 øyne

B: oddetall



$A \cup B$

$A \cap B$

I en befolkning av like mange menn og kvinner er 5% av mennene og 0.25% av kvinnene fargeblinde. En tilfeldig utvalgt person viser seg å være fargeblind. Hva er den *betingede* sannsynligheten for at vedkommende er en mann?

Eksperimentet: trekker en person og noterer biologisk kjønn / fargeblind

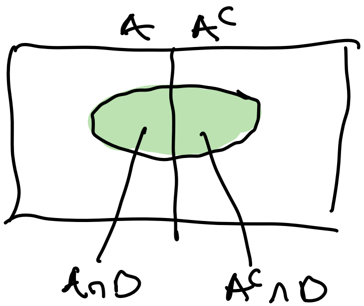
A: mann, A^c = kvinne

B: fargeblind, B^c = ikke fargeblind

$P(A) = 0.5 \Rightarrow P(A^c) = 0.5$

$P(B \mid A) = 0.05$, $P(B \mid A^c) = 0.0025$

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B \mid A) \cdot P(A)}{P(B)} = \underline{\underline{0.952}}$$



$$P(B) = \underbrace{P(A \cap B)}_{P(B \mid A) \cdot P(A)} + \underbrace{P(A^c \cap B)}_{P(B \mid A^c) \cdot P(A^c)}$$

$0.05 \cdot 0.5$ $\qquad$ $0.0025 \cdot 0.5$

Flere oppgaver: prøveeksamen oppg 1, K2021 1bc, H2020 1ab, se også oppgaver fra søsteremne under uke 10.

Mål: forstå hvordan dette brukes for diskrete stokastiske variabler

2.1 Urnemodeller

En urne inneholder n kuler. Antall mulige trekninger av r kuler er:

1. Ordnet utvalg, trekning med tilbakelegging

$$n^r$$

2. Ordnet utvalg, trekning uten tilbakelegging

$${}_nP_r = \frac{n!}{(n-r)!}$$

3. Ikke-ordnet utvalg, trekning uten tilbakelegging

$${}_nC_r = \binom{n}{r} = \frac{n!}{r!(n-r)!}$$

brukes som i binomisk fordeling

ellers ikke så aktuelt til eksamen

3 Stokastiske variabler

Mål: bygge forståelse!

3.1 Diskret

Kumulativ fordeling

$$F(x) = P(X \le x) = \sum_{t \le x} P(X = t)$$

Forventningsverdi

$$\mathrm{E}(X) = \mu_x = \sum_x x \cdot P(X = x)$$

Varians

$$\mathrm{Var}(X) = \sigma_x^2 = \mathrm{E}((X - \mu_x)^2) = \sum_x (x - \mu_x)^2 \cdot P(X = x)$$

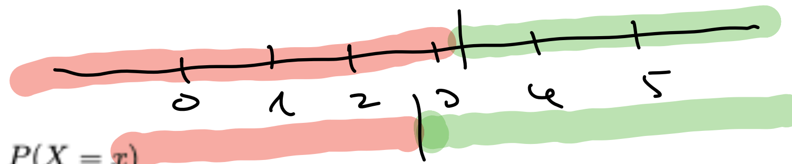
$$= \sum_x x^2 \cdot P(X = x) - \mu_x^2$$

Standardavvik

$$\mathrm{SD}(X) = \sigma_x = \sqrt{\mathrm{Var}(X)}$$

Husk at: $X \in \{0, 1, 2, 3, \dots\}$

$P(X > 3) = 1 - P(X \le 3)$



$P(X \ge 3) = 1 - P(X \le 2)$

Tegn tallinje og vær obs på om $P(X=3)$ skal være med i stykket

$$E(5 \cdot X) = \sum_x 5 \cdot x\, P(X=x)$$

$$= 5 \cdot E(X)$$

$$\mathrm{Var}(5X) = 25\, \mathrm{Var}(X)$$

Oppgaver: H2020 2ab

Mål: bygge forståelse!

3.2 Kontinuerlig

Kumulativ fordeling

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt$$

Forventningsverdi

$$\mathrm{E}(X) = \mu_x = \int_{-\infty}^{\infty} x \cdot f(x)\,dx$$

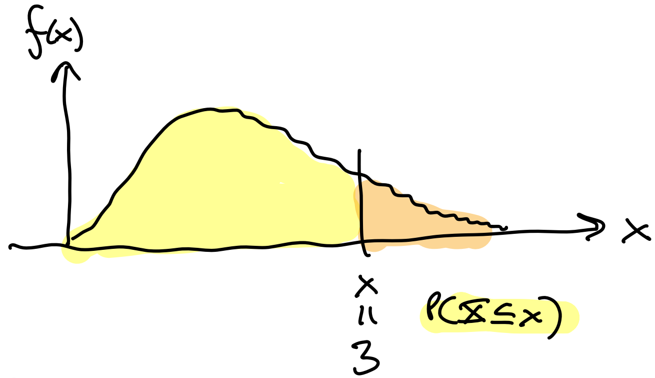
Varians

$$\mathrm{Var}(X) = \sigma_x^2 = \mathrm{E}((X-\mu_x)^2) = \int_{-\infty}^{\infty} (x-\mu_x)^2 \cdot f(x)\,dx$$

$$= \int_{-\infty}^{\infty} x^2 \cdot f(x)\,dx - \mu_x^2$$

Standardavvik

$$\mathrm{SD}(X) = \sigma_x = \sqrt{\mathrm{Var}(X)}$$



$P(X>3) = 1 - P(X \leq 3)$

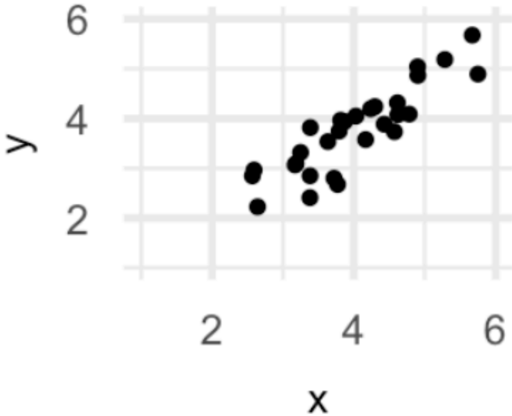
$P(X \geq 3) = P(X>3) = 1 - P(X \leq 3)$

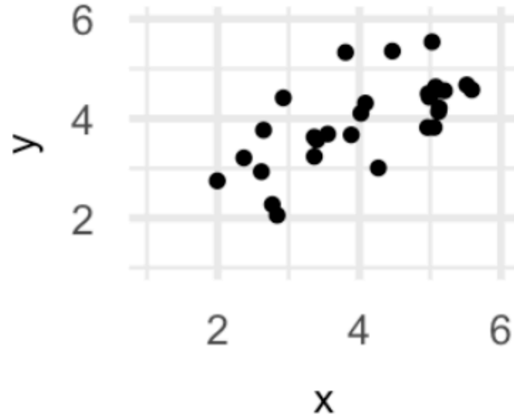
$P(X=3) = 0$

kont. stok.var

Oppgaver: ingen. → Vi skal ikke integrere på eksamen

Mål: bygge forståelse!





Forstå hva det betyr at

$\rho < 0$, $\rho > 0$

$\rho \approx 0$

3.3 Kovarians og korrelasjon

$\mathrm{Cov}(X,Y) = \mathrm{E}\left((X-\mu_x)\cdot(Y-\mu_y)\right) = \mathrm{E}(XY) - \mu_x\mu_y$

Korrelasjonskoeffisient: $\rho = \dfrac{\mathrm{Cov}(X,Y)}{\sigma_x \cdot \sigma_y}$

$\mathrm{Cov}(X,Y) = 0$ dersom X og Y er uavhengige.

8 Kovarians og korrelasjon

Estimator for **kovarians** $\mathrm{Cov}(X,Y)$ og **korrelasjon**:

$$\widehat{\mathrm{Cov}}(X,Y) = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y}),$$

$$R = \frac{\widehat{\mathrm{Cov}}(X,Y)}{S_x \cdot S_y},$$

der S_x og S_y er estimatorer for standardavvikene til X og Y.

Uke 1

Empirisk korrelasjon

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

Oppgaver: prøveeksamen 7a

3.4 Regneregler

ALLTID med på eksamen i et hvert grunnemne i statistikk!

$$P(a < X \leq b) = F(b) - F(a)$$

$$\mathrm{E}(aX + b) = a\mathrm{E}(X) + b$$

$$\mathrm{Var}(aX + b) = a^2\mathrm{Var}(X) \leftarrow \text{hvorfor:} \quad \sum (\overset{ax}{\downarrow}{x} - \mu)^2 P(X = x)$$

$$\mathrm{E}(aX + bY) = a\mathrm{E}(X) + b\mathrm{E}(Y)$$

$$\mathrm{Var}(aX + bY) = a^2\mathrm{Var}(X) + b^2\mathrm{Var}(Y) + 2ab\mathrm{Cov}(X, Y)$$

Vekten, X, til en tilfeldig valgt melkekartong på 1 liter har $E(X) = 1$ liter og Var$(X) = 0.01$ ~~liter~~ kg^2, og vekten til ulike kartonger er uavhengige. Hva er forventningsverdien og variansen til summen av vekten av to slike melkekartonger? $\quad X_1 + X_2$ vs $2X$

$$E(X_1 + X_2) = E(X_1) + E(X_2) = 1 + 1 = 2$$

$$\mathrm{Var}(X_1 + X_2) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \overbrace{2\,\mathrm{Cov}(X_1, X_2)}^{0 \text{ pga uavhengighet}}$$

$$= 2 \cdot \mathrm{Var}(X_1)$$

$$\mathrm{Var}(2X) = 4 \cdot \overset{0.01}{\overset{\parallel}{\mathrm{Var}(X)}}$$

Oppgaver: Prøveeksamen 4ab, K2021 2abc, 3b, H2020 3b + oppgaver om E og Var av estimatorer som prøveeksamen 8a.

VELDIG VIKTIG

$E(X_i) = \mu, \; \text{var}(X_i) = \sigma^2$

$\underbrace{\frac{1}{2}(X_1 + X_2)}_{\overline{X}}$

$\frac{1}{n}(X_1 + X_2 + \dots + X_n) = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}$

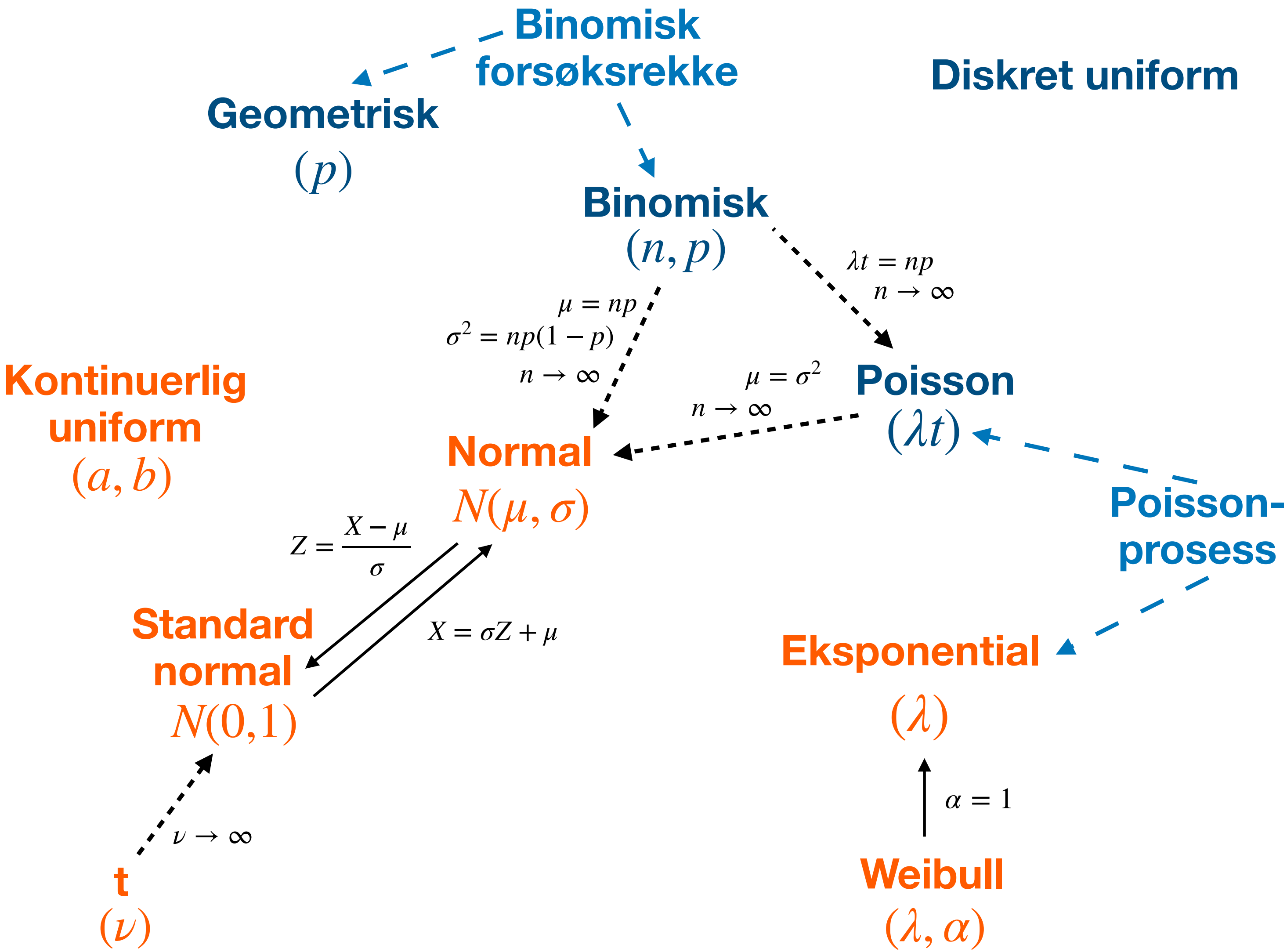
uavhengige X_i'er

$E(\overline{X}) = \mu$

$\text{Var}(\overline{X}) = \text{Var}\left(\frac{1}{n}\sum X_i\right) = \left(\frac{1}{n}\right)^2 \cdot \underbrace{\sum_{i=1}^{n} \underbrace{\text{Var}(X_i)}_{\sigma^2}}_{n \cdot \sigma^2}$

$= \frac{1}{n^2} \cdot n\sigma^2 = \frac{\sigma^2}{n}$

$SD(\overline{X}) = \frac{\sigma}{\sqrt{n}}$



Binomisk forsøksrekke
Diskret uniform
Geometrisk
(p)
Binomisk
(n, p)
$\lambda t = np$
$n \to \infty$
$\mu = np$
$\sigma^2 = np(1-p)$
$n \to \infty$
$\mu = \sigma^2$
$n \to \infty$
Poisson
(λt)
Kontinuerlig uniform
(a, b)
Normal
$N(\mu, \sigma)$
Poisson-prosess
$Z = \frac{X-\mu}{\sigma}$
$X = \sigma Z + \mu$
Standard normal
$N(0,1)$
Eksponential
(λ)
$\alpha = 1$
$\nu \to \infty$
t
(ν)
Weibull
(λ, α)

4 Sannsynlighetsfordelinger

Diskrete fordelinger: alltid minst en med på eksamen

Binomisk fordeling

$X \sim \text{Binom}(n, p)$ Hvor er nCr på din kalkulator?

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x},$$
$$\text{for } x = 0, 1, 2, \ldots, n$$
$$\text{E}(X) = np, \quad \text{Var}(X) = np(1-p)$$

Punktsannsynlighet $P(X=5)$

n, p / λt } oppgitt

↑

FORMLER

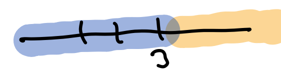
Geometrisk fordeling

$X \sim \text{Geom}(p)$

$$P(X = x) = p(1-p)^{x-1},$$
$$\text{for } x = 1, 2, \ldots$$
$$F(x) = P(X \leq x) = 1 - (1-p)^x$$
$$\text{E}(X) = \frac{1}{p}, \quad \text{Var}(X) = \frac{1-p}{p^2}$$

Hva er $P(X>3)$?

$= 1 - P(X \leq 3)$



binomisk tabell

geometrisk FORMEL se prøveeksamen

Poisson tabell

$\lambda t = 4$

Poissonfordeling

$X \sim \text{Poisson}(\lambda t)$

$$P(X = x) = \frac{(\lambda t)^x}{x!} e^{-\lambda t},$$
$$\text{for } x = 0, 1, 2, \ldots$$
$$\text{E}(X) = \lambda t, \quad \text{Var}(X) = \lambda t$$

$1 - P(X \leq 3) = 1 - 0{,}656$

$n = 6$

$p = 0.5$

$P(X>3)$

$= 1 - P(X \leq 3)$

$= 1 - 0.4335\ldots$

Oppgaver: Prøveeksamen 2, 9b, K2021 1a, 3a, se også oppgaver fra søsteremne under uke 10.

BINOMISK FORDELING: $P(X \leq x)$

		p											
n	x	0.05	0.1	0.15	0.2	0.25	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2	0	0.902	0.810	0.723	0.640	0.563	0.490	0.360	0.250	0.160	0.090	0.040	0.010
	1	0.998	0.990	0.978	0.960	0.938	0.910	0.840	0.750	0.640	0.510	0.360	0.190
3	0	0.857	0.729	0.614	0.512	0.422	0.343	0.216	0.125	0.064	0.027	0.008	0.001
	1	0.993	0.972	0.939	0.896	0.844	0.784	0.648	0.500	0.352	0.216	0.104	0.028
	2	1.000	0.999	0.997	0.992	0.984	0.973	0.936	0.875	0.784	0.657	0.488	0.271
4	0	0.815	0.656	0.522	0.410	0.316	0.240	0.130	0.062	0.026	0.008	0.002	0.000
	1	0.986	0.948	0.890	0.819	0.738	0.652	0.475	0.313	0.179	0.084	0.027	0.004
	2	1.000	0.996	0.988	0.973	0.949	0.916	0.821	0.688	0.525	0.348	0.181	0.052
	3	1.000	1.000	0.999	0.998	0.996	0.992	0.974	0.938	0.870	0.760	0.590	0.344
5	0	0.774	0.590	0.444	0.328	0.237	0.168	0.078	0.031	0.010	0.002	0.000	0.000
	1	0.977	0.919	0.835	0.737	0.633	0.528	0.337	0.187	0.087	0.031	0.007	0.000
	2	0.999	0.991	0.973	0.942	0.896	0.837	0.683	0.500	0.317	0.163	0.058	0.009
	3	1.000	1.000	0.998	0.993	0.984	0.969	0.913	0.812	0.663	0.472	0.263	0.081
	4	1.000	1.000	1.000	1.000	0.999	0.998	0.990	0.969	0.922	0.832	0.672	0.410
6	0	0.735	0.531	0.377	0.262	0.178	0.118	0.047	0.016	0.004	0.001	0.000	0.000
	1	0.967	0.886	0.776	0.655	0.534	0.420	0.233	0.109	0.041	0.011	0.002	0.000
	2	0.998	0.984	0.953	0.901	0.831	0.744	0.544	0.344	0.179	0.070	0.017	0.001
	3	1.000	0.999	0.994	0.983	0.962	0.930	0.821	0.656	0.456	0.256	0.099	0.016

POISSONFORDELING: $P(X_t \leq x)$

	λt						
x	0.1	0.5	1.0	1.5	2.0	2.5	3.0
0	0.9048	0.6065	0.3679	0.2231	0.1353	0.0821	0.0498
1	0.9953	0.9098	0.7358	0.5578	0.4060	0.2873	0.1991
2	0.9998	0.9856	0.9197	0.8088	0.6767	0.5438	0.4232
3	1.0000	0.9982	0.9810	0.9344	0.8571	0.7576	0.6472
4	1.0000	0.9998	0.9963	0.9814	0.9473	0.8912	0.8153
5	1.0000	1.0000	0.9994	0.9955	0.9834	0.9580	0.9161
6	1.0000	1.0000	0.9999	0.9991	0.9955	0.9858	0.9665
7	1.0000	1.0000	1.0000	0.9998	0.9989	0.9958	0.9881
8	1.0000	1.0000	1.0000	1.0000	0.9998	0.9989	0.9962
9	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9989
10	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997
11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999
12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

	λt						
x	3.5	4.0	4.5	5.0	5.5	6.0	6.5
0	0.0302	0.0183	0.0111	0.0067	0.0041	0.0025	0.0015
1	0.1359	0.0916	0.0611	0.0404	0.0266	0.0174	0.0113
2	0.3208	0.2381	0.1736	0.1247	0.0884	0.0620	0.0430
3	0.5366	0.4335	0.3423	0.2650	0.2017	0.1512	0.1118
4	0.7254	0.6288	0.5321	0.4405	0.3575	0.2851	0.2237
5	0.8576	0.7851	0.7029	0.6160	0.5289	0.4457	0.3690
6	0.9347	0.8893	0.8311	0.7622	0.6860	0.6063	0.5265
7	0.9733	0.9489	0.9134	0.8666	0.8095	0.7440	0.6728
8	0.9901	0.9786	0.9597	0.9319	0.8944	0.8472	0.7916
9	0.9967	0.9919	0.9829	0.9682	0.9462	0.9161	0.8774
10	0.9990	0.9972	0.9933	0.9863	0.9747	0.9574	0.9332

$$P(\underbrace{X>6}_{A} \mid \underbrace{X>3}_{B}) = \frac{P(X>6 \cap X>3)}{P(X>3)} = \frac{P(X>6)}{P(X>3)}$$

Husk:

minst en
en eller flere feil } $X \geq 1$

0, 1, 2 . . .

$$P(X \geq 1) = 1 - P(X=0)$$

NB ordbok er lov

Kontinuerlige fordelinger

Eksponentialfordeling

$$T \sim \text{Eksponential}(\lambda)$$
$$f(t) = \lambda e^{-\lambda t}, \quad \text{for } t > 0$$
$$F(t) = P(T \leq t) = 1 - e^{-\lambda t}$$
$$R(t) = P(T > t) = 1 - F(t) = e^{-\lambda t}$$
$$\text{E}(T) = \frac{1}{\lambda}, \quad \text{Var}(T) = \frac{1}{\lambda^2}$$

Weibullfordeling

$$T \sim \text{Weibull}(\alpha, \lambda)$$
$$f(t) = \alpha\lambda(\lambda t)^{\alpha-1} e^{-(\lambda t)^\alpha}, \quad \text{for } t > 0$$
$$F(t) = P(T \leq t) = 1 - e^{-(\lambda t)^\alpha}$$
$$R(t) = P(T > t) = 1 - F(t) = e^{-(\lambda t)^\alpha}$$
$$\cancel{\text{E}(T) = \frac{1}{\lambda}\Gamma(1 + \tfrac{1}{\alpha})}$$
$$\cancel{\text{Var}(T) = \frac{1}{\lambda^2}\left(\Gamma(1 + \tfrac{2}{\alpha}) - \Gamma(1 + \tfrac{1}{\alpha})^2\right)}$$

ikke på eksamen

(kalkulator har ikke "Γ")

levetid av komponenter

$$T \sim \text{eksp}(\lambda = 3)$$

$$P(T > 2) = e^{-\overset{3}{\lambda} \cdot \overset{2}{t}} = \underline{e^{-6}}$$

$$\Vert$$

$$1 - P(T \leq 2)$$

tilsvarende for Weibull

Oppgaver: Prøveeksamen 9ab, H2020 1c,

4.2 Begreper fra levetidsanalyse

Pålitelighetsfunksjon

$$R(t) = 1 - F(t) = P(X > t)$$

Sviktrate

$$h(t) = \frac{f(t)}{R(t)}$$

Systempålitelighet seriekobling ← juletrelys, solceller 

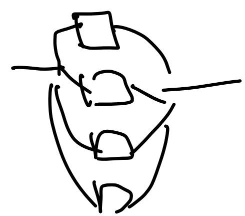
alle må fungere hvis systemet skal fungere

$$R(t) = R_1(t) \cdot R_2(t) \cdots R_n(t)$$

Systempålitelighet parallellkobling

høyttaler

$$R(t) = 1 - F_1(t) \cdot F_2(t) \cdots F_n(t)$$



nok at en fungerer

Oppgaver: Prøveeksamen 9c.

Normalfordeling

$$X \sim \mathrm{N}(\mu, \sigma)$$

$$f(x) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad \text{for } -\infty < x < \infty$$

$$F(x) = P(X \leq x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

$$\mathrm{E}(X) = \mu, \quad \mathrm{Var}(X) = \sigma^2$$

Standard normalfordeling

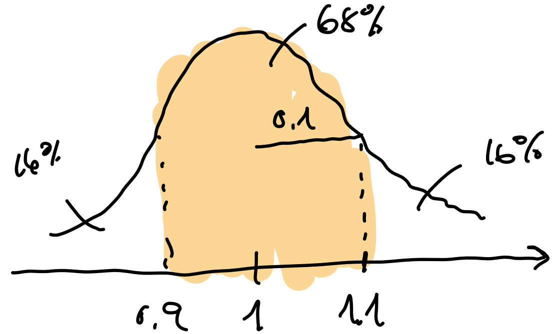
$$Z \sim \mathrm{N}(0, 1)$$

$$f(z) = \phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, \quad \text{for } -\infty < z < \infty$$

$$F(z) = P(Z \leq z) = \Phi(z)$$

$$\mathrm{E}(Z) = 0, \quad \mathrm{Var}(Z) = 1$$

Den viktigste fordelingen i emnet: HELT sikkert på eksamen



$X \sim N(1, 0.1)$

En fabrikk masseproduserer en artikkel. Vekten av én artikkel antas å være normalfordelt med forventningsverdi 1kg og standardavvik 0.1 kg.

Hva er sannsynligheten for at en artikkel veier mer enn 1.1 kg?

$$P(X > 1.1) = P\left(\frac{X - \overset{1}{\mu}}{\underset{0.1}{\sigma}} > \frac{1.1 - 1}{0.1}\right) = P(Z > 1)$$

$$= 1 - P(Z \leq 1)$$

$$= 1 - \phi(1) = 1 - 0.8413 = \underline{\underline{0.1587}}$$

STANDARD NORMALFORDELING: $\Phi(z) = P(Z \leq z)$

z	...0	...1	...2	...3	...4	...5	...6	...7	...8	...9
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545

EKSAMEN

- hvor DU skal sitte er NÅ klart på Stadweb!

- Noen er plassert der det er PC - og slipper da å ta med egen laptop

- HUSK: brukernavn & passord

  det du logger på Inspera med!

  (hvis du sitter på NTNU PC kan du jo ikke bruke autolagret brukernavn & passord)

Spørsmål fra chat:

→ Kalkulator: husk nye batterier. Det er ikke kalkulator på PC'en.

→ På studweb skal det rtig eksamenshlr stå om du har fått NTNU PC. Er du i tvil ta kontakt med Eksamenskontoret ← det er de som arrangerer eksamen!

→ Verdier for økonomiske ... er ikke er i tabell: DET kommer ikke på eksamen. Alle tall er nøyaktig valgt ut.

→ Innlogging i Inspera: bare brukernavn + passord (sier eksamenskontoret)

ikke fødselsnummer sier

→ Prøven er ikke aktivert i Inspera ennå, trolig torsdag – men vi må ha kommet i havn med de siste øvingsgodkjenningene

NTNU i Trondheim

E-post: eksamen@adm.ntnu.no

Åpningstid for telefonhenvendelser:

Mandag–fredag 09.00–11.30

Telefon: 73 59 66 00

Besøkshenvendelser:

Besøksadresse: G470, Gunnerusgate 1, 4. etasje, Kalvskinnet

Åpent etter avtale. Skal du levere noe, benytt postkassen utenfor døren.

Postadresse: Eksamenskontoret NTNU, Postboks 8900, Torgarden, 7491 Trondheim

NTNU i Gjøvik - Studenttorget

E-post: eksamen@gjovik.ntnu.no

Telefon: 61 13 54 00

Besøksadresse: Studenttorget i Gneis-bygget, NTNU i Gjøvik, Teknologiveien 22, 2815 Gjøvik

Postadresse: NTNU i Gjøvik, postboks 191, 2802 Gjøvik

NTNU i Ålesund

E-post: studier@alesund.ntnu.no

Telefon: 73 59 50 00

Postadresse: NTNU i Ålesund, postboks 1517, 6025 Ålesund

KONTAKTINFO EKSAMENSKONTORET

(4.1) Regneregler normalfordeling

Vi ser på n uavhengige stokastiske variabler $X_1, X_2, \dots, X_n$ slik at $X_i \sim N(\mu, \sigma)$, for $i = 1, \dots, n$. Da er

$$\bar{X} = \frac{X_1 + X_2 + \cdots + X_n}{n} \sim \mathrm{N}\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

E, Var: kunne hente fra (3.4)

$n = 3$

$$E\left(\tfrac{1}{3}X_1 + \tfrac{1}{3}X_2 + \tfrac{1}{3}X_3\right) = \tfrac{1}{3}\mu + \tfrac{1}{3}\mu + \tfrac{1}{3}\mu = \mu$$

$$\mathrm{Var}\left(\tfrac{1}{3}X_1 + \tfrac{1}{3}X_2 + \tfrac{1}{3}X_3\right) = \tfrac{1}{9}\sigma^2 + \tfrac{1}{9}\sigma^2 + \tfrac{1}{9}\sigma^2$$

$$= \tfrac{3}{9}\sigma^2 = \tfrac{1}{3}\sigma^2$$

og

$$X_1 + X_2 + \cdots + X_n \sim \mathrm{N}\left(n\mu, \sqrt{n} \cdot \sigma\right)$$

4.1.1 Normaltilnærminger

Binom$(n, p) \approx \mathrm{N}(np, \sqrt{np(1-p)})$ hvis $np(1-p) \geq 5$,
Poisson$(\lambda t) \approx \mathrm{N}(\lambda t, \sqrt{\lambda t})$ hvis $\lambda t > 10$

4.1.2 Sentralgrenseteoremet

Dersom $X_1, X_2, \dots, X_n$ er uavhengige stokastiske variabler med forventning $\mathrm{E}(X_i) = \mu$ og varians $\mathrm{Var}(X_i) = \sigma^2$, for $i = 1, \dots, n$, og dersom $n > 30$, så er

$$\bar{X} = \frac{X_1 + X_2 + \cdots + X_n}{n} \text{ tilnærmet } \mathrm{N}\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

og

$$X_1 + X_2 + \cdots + X_n \text{ tilnærmet } \mathrm{N}\left(n\mu, \sqrt{n} \cdot \sigma\right)$$

Vekten av én fabrikkprodusert artikkel antas å være normalfordelt med forventningsverdi 1kg og standardavvik 0.1 kg.

0.1587

Vi studerer 10 artikler. Hva er sannsynligheten for at gjennomsnittlig vekt av 10 artikler er høyere enn 1.1 kg?

Oppgaver om normalfordelingn: Prøveeksamen 3, H2020 4, se også oppgaver fra søsteremne under uke 10.

Vekten av én fabrikkprodusert artikkel antas å være normalfordelt med forventningsverdi 1kg og standardavvik 0.1 kg.

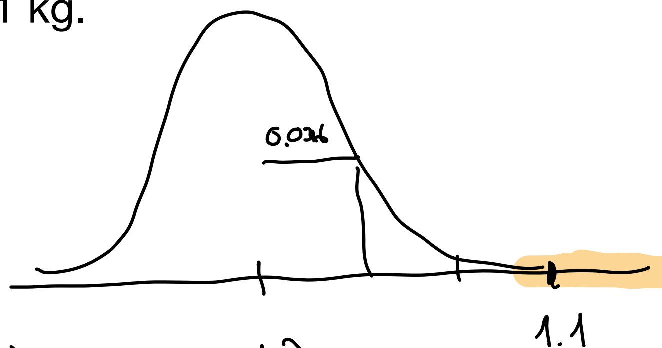
Vi studerer 10 artikler. Hva er sannsynligheten for at gjennomsnittlig vekt av 10 artikler er høyere enn 1.1 kg.

er

$$\bar{X} = \frac{X_1 + X_2 + \cdots + X_n}{n} \sim \mathrm{N}\left(\mu, \underbrace{\frac{\sigma}{\sqrt{n}}}_{SD(\bar{X})}\right)$$

$$Var(\bar{X}) = \frac{\sigma^2}{n}$$

$$\bar{X} \sim N\left(1, \underbrace{\frac{0.1}{\sqrt{10}}}_{0.0316}\right)$$



$$P(\bar{X} > 1.1) = P\left(\underbrace{\frac{\bar{X} - E(\bar{X})}{SD(\bar{X})}}_{\text{standardiser}} > \underbrace{\frac{1.1 - 1}{0.0316}}_{3.16}\right) = 1 - \underbrace{\Phi(3.16)}$$

$$= 1 - 0.9992 = 0.0008$$

STANDARD NORMALFORDELING: $\Phi(z) = P(Z \leq z)$

z	...0	...1	...2	...3	...4	...5	...6	...7	...8	...9
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999

Så langt!

Eksamen: 60-70% er sannsynlighetsregning

Går du for C-E? Det er mulig med å i hovedsak svare på spørsmål fra sannsynlighetsregning.

A-B: Da må også statistisk inferens være på plass!

E, Var

5 Punktestimering

5.1 Forventningsverdi og varians

For et tilfeldig utvalg $X_1, X_2, \ldots, X_n$ der $\mathrm{E}(X_i) = \mu$ og $\mathrm{Var}(X_i) = \sigma^2$, for $i = 1, \ldots, n$, så er

Estimator for forventningsverdien (μ):

$$\hat{\mu} = \bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

Estimator for varians (σ^2):

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

Estimator for standardavviket (σ):

$$S = \sqrt{S^2}$$

5.2 Sannsynligheten p i binomisk fordeling

Dersom X teller antall suksesser i en binomisk forsøksrekke av n forsøk så er estimator for sannsynligheten for suksess (p):

$$\hat{p} = \frac{X}{n}$$

5.3 Raten λ i poissonfordelingen

Dersom X teller antall hendelser i en poissonprosess med rate λ over et intervall/område av lengde/størrelse t, så er estimator for raten (λ):

$$\hat{\lambda} = \frac{X}{t}$$

uttrykk for 4 estimatorer

$\underbrace{\mu \qquad \sigma}_{N} \qquad p \qquad \lambda$

bin n — Poisson t

Oppgaver: Prøveeksamen 8, se også oppgaver fra søsteremne under uke 10.

La $X_1, X_2, \ldots, X_n$ være uavhengige og normalfordelte $N(\mu, \sigma)$ stokastiske variabler som angir mengden fosfor i prøver fra et renseanlegg (som prøveeksamen oppgave 6). Vi vil estimere μ.

En naturlig estimator er $\bar{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$ og den har vi regnet på mange ganger og vet at E$(\bar{X}) = \mu$ og SD$(\bar{X}) = \dfrac{\sigma}{\sqrt{n}}$.

La $n = 3$. En konkurrerende estimator er $\mu^* = \dfrac{1}{4}(X_1 + 2X_2 + X_3)$.

Hvilken estimator for μ er best? μ^* eller $\bar{X}$?

1) Forventningsrett: $E(\bar{X}) = \underline{\mu}$

$$E(\mu^*) = E\left(\frac{1}{4}(X_1 + 2X_2 + X_3)\right) = \frac{1}{4}\left[\overset{\mu}{E(X_1)} + 2\overset{\mu}{E(X_2)} + \overset{\mu}{E(X_3)}\right] = \frac{1}{4}\cdot 4\mu = \underline{\underline{\mu}}$$

2) Minst varians:

$$\mathrm{Var}(\bar{X}) = \frac{\sigma^2}{3} \qquad SD(\bar{X}) = \frac{\sigma}{\sqrt{3}}$$

$$\mathrm{Var}(\mu^*) = \left(\frac{1}{4}\right)^2 \left[\mathrm{Var}(X_1) + 2^2 \cdot \mathrm{Var}(X_2) + \mathrm{Var}(X_3)\right]$$

$$= \frac{1}{4^2}\left[6 \cdot \sigma^2\right] = \frac{6}{16} \cdot \sigma^2 = \frac{3}{8}\sigma^2$$

$$\frac{\sigma^2}{3} \quad \text{vs} \quad \frac{3}{8}\sigma^2$$

$\bar{X}$ $\qquad$ μ^*

$0.33 \cdot \sigma^2$ $\qquad$ $0.375 \cdot \sigma^2$

<u>velg denne</u>

6 Konfidensintervall

6.1 Forventningsverdi μ

For et tilfeldig utvalg $X_1, X_2, \ldots, X_n$, $X_i \sim N(\mu, \sigma)$, $i = 1, \ldots, n$, der standardavviket σ er *kjent*, så er

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

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

Intervall som vi har stor tiltro til at denn paremeter ligger innenfor.

Dersom standardavviket σ er *ukjent*, så er

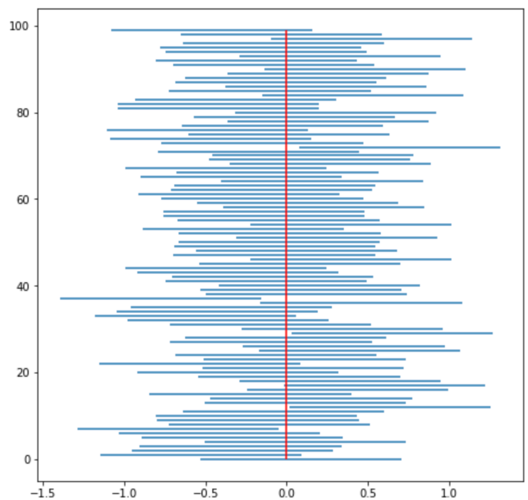
$$\left[\bar{X} - t_{\alpha/2,n-1} \cdot \frac{S}{\sqrt{n}},\ \ \bar{X} + t_{\alpha/2,n-1} \cdot \frac{S}{\sqrt{n}}\right]$$

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

På eksamen $\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$

Oppgir $\underbrace{\sum_{i=1}^{n}(X_i-\bar{X})^2}_{8}$

$S = \sqrt{\frac{1}{n-1}}$



Oppgaver: Prøveeksamen 5, uke 10 søsteremne

Kritiske verdier i standard normalfordelingen

$$P(Z > z_\alpha) = \alpha$$

α	z_α
0.20	0.842
0.10	1.282
0.05	1.645
0.025	1.960
0.02	2.054
0.01	2.326
0.005	2.576
0.001	3.090
0.0005	3.291
0.0001	3.719

95% KI

$\alpha = 0.05$

$\frac{\alpha}{2} = 0.025$

$z_{0.025} = 1.96$

$n=10, \quad \frac{\alpha}{2} = 0.025$

$n-1$

$t_{0.025, 9} = 2.262$

Kritiske verdier i t-fordelingen.

$$P(T > t_{\alpha,\nu}) = \alpha$$

	α							
ν	0.2	0.1	0.05	0.025	0.01	0.005	0.001	0.0005
1	1.376	3.078	6.314	12.706	31.821	63.657	318.309	636.619
2	1.061	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.978	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.941	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.920	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.906	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.896	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.889	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.883	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.879	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.876	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.873	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.870	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.868	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.866	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.865	1.337	1.746	2.120	2.583	2.921	3.686	4.015

6.2 Sannsynligheten p i binomisk fordeling

Under forutsetningen om at $n\hat{p}(1-\hat{p}) \geq 5$, så er

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

et tilnærmet $100(1-\alpha)\%$ konfidensintervall for p.

6.3 Raten λ i poissonfordelingen

Under forutsetningen om at $\hat{\lambda}t > 10$, så er

$$\left[\hat{\lambda} - z_{\alpha/2} \cdot \sqrt{\frac{\hat{\lambda}}{t}},\ \hat{\lambda} + z_{\alpha/2} \cdot \sqrt{\frac{\hat{\lambda}}{t}}\right]$$

et tilnærmet $100(1-\alpha)\%$ konfidensintervall for λ.

7 Hypotesetesting

Noen begreper:

- *Type I-feil:* Forkaste nullhypotesen H_0 selv om H_0 er sann.
- *Type II-feil (eller type 2-feil):* Ikke forkaste nullhypotesen H_0 selv om den alternative hypotesen H_1 er sann.
- *Teststyrke:* Teststyrken til en test er sannsynligheten for å forkaste nullhypotesen H_0 til fordel for den alternative hypotesen H_1 når den alternative hypotesen er sann og vi kjenner den riktige parameterverdien.
- *P-verdi:* P-verdien er sannsynligheten for det vi har observert, eller noe mer ekstremt i retning den alternative hypotesen H_1, når vi antar at nullhypotesen H_0 er sann.

α = signifikanssannsynlighet

= det vi tåler av type I-feil

:= P(type I-feil)

	Ikke forkaste H_0	Forkaste H_0
H_0 sann	Korrekt	Type I feil
H_0 falsk	Type II feil	Korrekt

NOTATER FOR HYPOTESETEST er LAGT TIL etter timen

7.1 Forventningsverdi μ

1) Fra spørsmålet velge H_0: ← =

H_1: ← spørsmålet

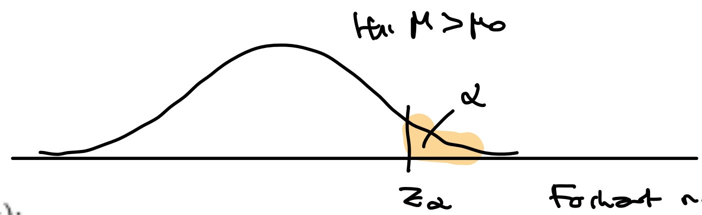
μ, p

Testobservator for $H_0 : \mu = \mu_0$ mot

1. $H_1 : \mu \neq \mu_0$, eller
2. $H_1 : \mu > \mu_0$, eller → forkast for store verdier
3. $H_1 : \mu < \mu_0$, → forkast for små verdier

dersom standardavviket σ er *kjent* (Z-test):

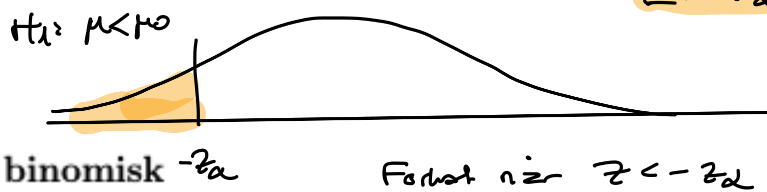
$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \overset{H_0}{\sim} N(0,1)$$



Forkast når $Z > z_\alpha$

og dersom standardavviket er *ukjent* (T-test):

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \overset{H_0}{\sim} t_{n-1}$$

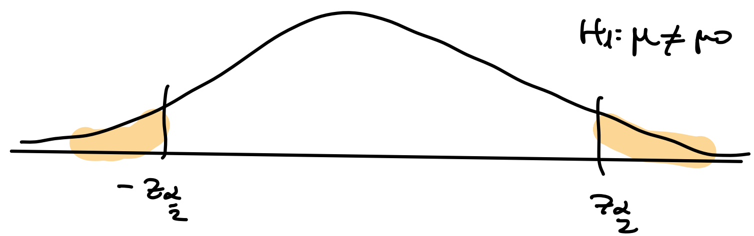


Forkast når $Z < -z_\alpha$

7.2 Sannsynligheten p i binomisk fordeling

Testobservator for $H_0 : p = p_0$ (Z-test):

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \overset{H_0}{\approx} N(0,1)$$



Forkast når $Z < -z_{\frac{\alpha}{2}}$ eller $Z > z_{\frac{\alpha}{2}}$

Oppgaver: Prøveeksamen 6, K2021 4, H2020 3b.

Systolisk blodtrykk: Vi vil teste om gjennomsnittlig systolisk blodtrykk hos en populasjon av kvinner med en spesiell sykdom er høyere enn 120mmHg. La μ være det (ukjente) gjennomsnittlige blodtrykket i denne populasjonen.

$H_0: \mu = 120$ mot $H_1: \mu > 120$

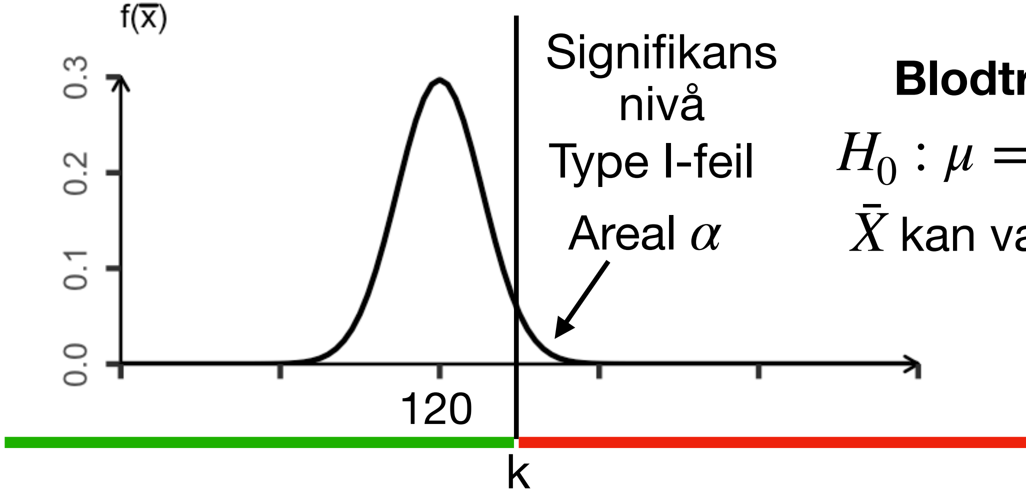
La $X_1, X_2, \ldots, X_n$ være blodtrykket til n tilfeldig valgte kvinner fra populasjonen med den spesielle sykdommen og anta at X_i er normalfordelt med forventningsverdi $\mathrm{E}(X_i) = \mu$ og $\mathrm{SD}(X_i) = \sigma$.

Hva kan vi bruke som testobservator? $\bar{X} \sim N(\mu_0, \frac{\sigma}{\sqrt{n}})$

når H_0 er sann

Lettest å heller tenke at vi bruker

$$Z = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1) \text{ når } H_0 \text{ er sann.}$$



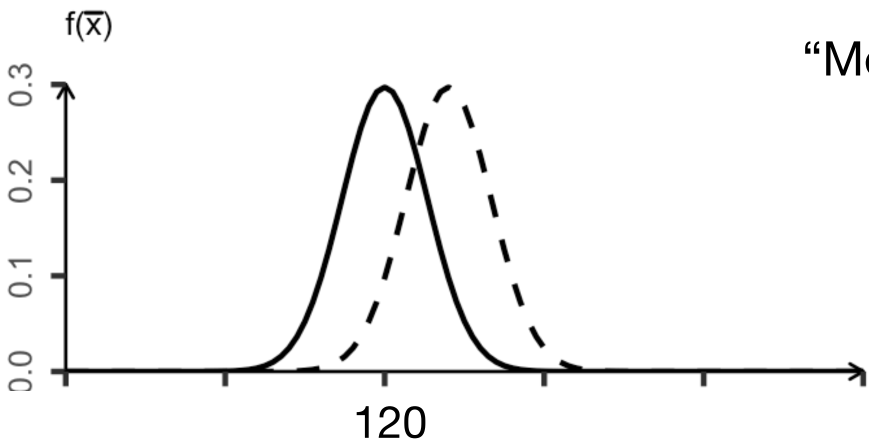
Blodtrykk-eksemplet

$H_0 : \mu = 120$ **mot** $H_1 : \mu > 120$

$\bar{X}$ kan være testobservator for μ

$k = \overset{120}{\mu_0} + z_\alpha \cdot \frac{\sigma}{\sqrt{n}}$

Fordeling til testobservator når nullhypotesen er sann



"Mer generelt" $Z = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$

k blir da kritisk verdi i standard normalfordelingen

$Z > z_\alpha$

.. sammen med fordelingen for en gitt alternativ hypotese

Hvis vi velger signifikansnivå 5%, når forkaster vi nullhypotesen?

Forkast når $z > z_\alpha$

regn ut

slå opp

«samme» retning som H_1!

høyresidig test

$z_{0.05} = 1.645$

Vi har observert data for $n = 30$ kvinner og regnet ut at $\bar{x} = 121.6$. Videre er det kjent at $\sigma = 8.8$. Vil vi forkaste nullhypotesen? Hva er *p*-verdien?

- *P-verdi:* *P*-verdien er sannsynligheten for det vi har observert, eller noe mer ekstremt i retning den alternative hypotesen H_1, når vi antar at nullhypotesen H_0 er sann.

$$z = \frac{121.6 - 120}{\underbrace{\frac{8.8}{\sqrt{30}}}_{1.61}} = 0.9959 \approx 1.00$$

$P(Z > 1.00) = 1 - \Phi(1.00) = 1 - 0.8413 = 0.1587 =$ ganske sannsynlig

$\Rightarrow$ $0.1587 > 0.05$ $\Rightarrow$ ikke forkast H_0

Hva er teststyrken til testen når $\mu = 124$? ← Skille mellom karakter A og B

- *Teststyrke:* Teststyrken til en test er sannsynligheten for å forkaste nullhypotesen H_0 til fordel for den alternative hypotesen H_1 når den alternative hypotesen er sann og vi kjenner den riktige parameterverdien.

P(forkaste H_0 når H_1 er sann og spesielt $\mu = 124$)

$Z > 1.645$

$$\frac{\bar{X} - 120}{\frac{8.8}{\sqrt{30}}} > 1.645$$

$$\bar{X} > \underbrace{120}_{\mu_0} + \underbrace{1.645}_{z_\alpha} \cdot \underbrace{\overbrace{\left(\frac{8.8}{\sqrt{30}}\right)}^{1.61}}_{\frac{\sigma}{\sqrt{n}}} = 122.6429$$

$$= P(\bar{X} > 122.6429 \text{ gitt } E(\bar{X}) = 124)$$

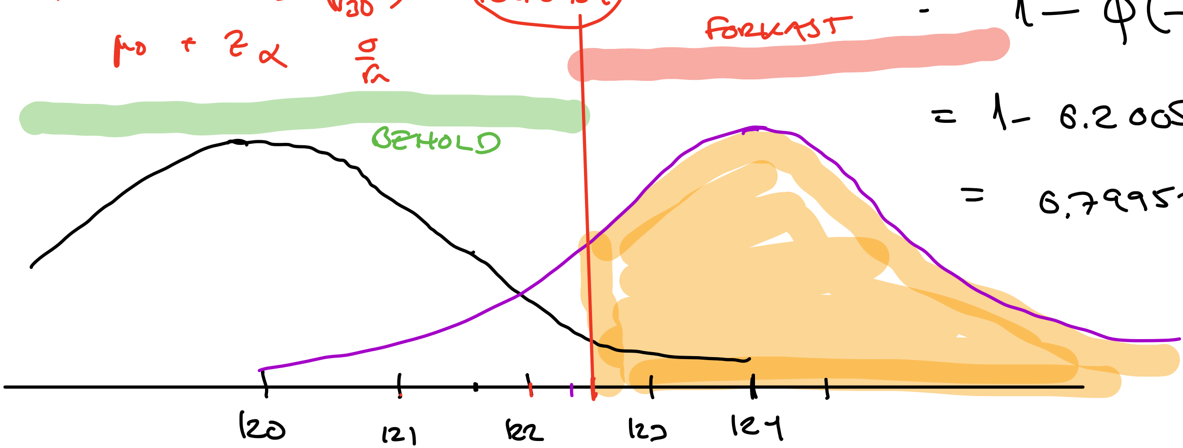
$$= P\left(\frac{\bar{X} - 124}{1.61} > \frac{122.6429 - 124}{1.61}\right)$$

$$= 1 - \Phi(-0.84)$$

$$= 1 - 0.2005$$

$$= 0.7995 \approx \underline{\underline{0.80}}$$

≠ ok tall



STANDARD NORMALFORDELING: $\Phi(z) = P(Z \leq z)$

z	..0	..1	..2	..3	..4	..5	..6	..7	..8	..9
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621

STANDARD NORMALFORDELING: $\Phi(z) = P(Z \leq z)$

z	..0	..1	..2	..3	..4	..5	..6	..7	..8	..9
-0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
-0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
-0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
-0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
-0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
-0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
-0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
-0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
-0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
-0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641

9 Enkel lineær regresjon

La $(x_1, Y_1), (x_2, Y_2), \ldots, (x_n, Y_n)$ være n uavhengige par der x-ene er kjente tall, og Y-ene er stokastiske variabler slik at

$$Y_i | X_i = x_i \sim \mathrm{N}(\beta_0 + \beta_1 x_i, \sigma) \text{ for } i = 1, \ldots, n$$

En annen måte å formulere modellen på er:

$$Y_i = \beta_0 + \beta_1 x_i + e_i, \quad e_i \sim \mathrm{N}(0, \sigma)$$
$$\text{for } i = 1, \ldots, n$$

Minste kvadratsums estimatorer for β_0 og β_1:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}, \quad \text{og}$$
$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{x}.$$

Estimert regresjonslinje:

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$$

Minste kvadratsums rette linje $y = a + bx$

$$b = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$
$$a = \bar{y} - b\bar{x}.$$

Oppgaver: Prøveeksamen 7, H2020 4 (ikke c)

Estimator for varians σ^2 er:

$$S^2 = \frac{1}{n-2}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$$
$$= \frac{1}{n-2}\sum_{i=1}^{n}(Y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2$$

og estimator for standardavviket σ, er $S = \sqrt{S^2}$.

Godhetsmål for regresjonsmodellen:

$$R^2 = \frac{\text{SST} - \text{SSE}}{\text{SST}}, \quad \text{der}$$

$$\text{SST} = \sum_{i=1}^{n}(Y_i - \bar{Y})^2 \text{ og,}$$

$$\text{SSE} = \sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 = \sum_{i=1}^{n}(Y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2$$

Konfidensintervall for β_1:

$$\left[\hat{\beta}_1 - t_{\alpha/2,n-2} \cdot \text{SE}(\hat{\beta}_1), \;\; \hat{\beta}_1 + t_{\alpha/2,n-2} \cdot \text{SE}(\hat{\beta}_1)\right]$$

der

$$\text{SE}(\hat{\beta}_1) = \frac{S}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}}$$

Testobservator for $H_0 : \beta_1 = 0$ mot $H_0 : \beta_1 \neq 0$:

$$T = \frac{\hat{\beta}_1 - 0}{\text{SE}(\hat{\beta}_1)} \overset{H_0}{\sim} t_{n-2}$$

1001 , 1003

Ikke eksamen!

ISTA/G/T 1001/2/3
Statistikk for ingeniører

Informasjon

Informasjon om emnet

Arbeidskrav, pensum og ressurser

Referansegruppe og fagteam

Digital plenumstime

ISTA: Campus Ålesund

ISTG: Campus Gjøvik

ISTT1001/3: Campus Trondheim

ISTT1002: Campus Trondheim

Undervisning og øvinger

Fellesmodul

STACK-øvinger

Python/Jupyter

Digitalt forum

Prosjektgrupper: påmelding og kontrakt

Eksamen

Informasjon om eksamen

Formelark og tabeller

Tidligere eksamensoppgaver

Eksamensforberedelser og prøveeksamen

Fellesmodul: Uke 10

ISTA/G/T 1001/2/3
Statistikk for ingeniører

Fellesmodul: Uke 10

Introduksjon

Denne uke skal du jobbe mot eksamen fredag 29.10 kl 9.00-12.00!

Ukas aktiviteter:

1. **Mest relevant er det å jobbe med prøveeksamen (se under), dernest eksamen fra konten 2021 og høsten 2020 (som du finner under Tidligere eksamensoppgaver) og dernest STACK-oppgavene i år, og deretter eldre eksamensoppgaver.**
2. Du har *muligheten* for å starte uka med å følge den digitale plenumsforelesningen som zoom-webinar på mandag 25.oktober kl 14.15-16.00 (se info under - samme lenke hele semestret). Etter avtale med referansegruppa vil Mette gå gjennom formelarket (og se kort på tabelloppslag) og forklare og regne eksempler. Dere kan stille spørsmål underveis i chat, og det blir som vanlig gjort opptak (ikke alle vil ønske å delta og ikke alle har tid den ekstra timen 15.15-16).
3. Du kan se repetisjonsvideoer fra hele fellesdelen (minus det nye om Weibull og pålitelighet i uke 5). Disse finner du i Panopto, se lenker under.
4. Du kan repetere alle STACK-øvingene ved å se på "Repetisjonssettene": stack.math.ntnu.no
5. Du kan selvfølgelig se på Jupyter-notatbøker, men de kan du ikke bruke på eksamen. Du *kan* få utskrift fra Python på eksamen - s.ntnu.no/isthub. Se oppgave 7 i Prøveeksamen som et eksempel på det.
6. Du får hjelp med eksamensspørsmål på på din campus (se under campussidene) eller ved å stille spørsmål i forumet under ekssamensspørsmål:https://mattelab.math.ntnu.no/c/istx100y/eksamenssporsmal/160 eller av Mette på zoom (lenke for påmelding under).
7. For de som ikke har fått godkjent 6 av 9 øvinger er aller siste frist STACK øving 9 onsdag 28.10 kl 23.59. Alle som mangler øvinger har fått henvendelse fra Eksamenskontoret eller fra Siebe og Siebe er kontaktperson ved spørsmål om dette.
8. **Fredag 29.10 møter du opp på anvist sted - det er klart 3 dager før på StudWeb. Husk å være der i god tid før kl 9.00 og ha med PC, kabler, kalkulator, skrivesaker, mat/drikke - og gyldig studentlegitimasjon!**

https://innsida.ntnu.no/wiki/-/wiki/Norsk/Regler+for+eksamen
https://innsida.ntnu.no/wiki/-/wiki/Norsk/Digital+skoleeksamen+-+for+studenter

Informasjon om eksamen
Formelark og tabeller
Tidligere eksamensoppgaver
Eksamensforberedelser og prøveeksamen

ISTA/G/T 1001/2/3
Statistikk for ingeniører

Informasjon

Fellesmodul: Uke 10

Repetisjonsvideo

Her er en repetisjonsvideo med Thea, som er hun som har laget alle temavideoene. Videoen ble laget i 2020, men er like relevant i 2021!

Video: oversiktsforelesning

Slides: oversiktsforelesning.pdf

Notater: notateroppsummering.pdf

Veiledning på zoom tirsdag 26.10 og onsdag 27.10

Har du faglige spørsmål og vil møte koordinator Mette på zoom for veiledning i uke 10?

Meld deg på 15-minutters møter her for tirsdag 26.10 kl 13.15-16 og onsdag 27.10 kl 13.15-16.
Påmelding zoom-veiledning med Mette

Eksamen

Informasjon om eksamen

Formelark og tabeller

Tidligere eksamensoppgaver

Eksamensforberedelser og prøveeksamen

ISTA/G/T 1001/2/3
Statistikk for ingeniører

Informasjon

Informasjon om emnet

Arbeidskrav, pensum og ressurser

Referansegruppe og fagteam

Digital plenumstime

ISTA: Campus Ålesund

ISTG: Campus Gjøvik

ISTT1001/3: Campus Trondheim

ISTT1002: Campus Trondheim

Undervisning og øvinger

Fellesmodul

STACK-øvinger

Python/Jupyter

Digitalt forum

Prosjektgrupper: påmelding og kontrakt

Eksamen

Informasjon om eksamen

Formelark og tabeller

Tidligere eksamensoppgaver

Eksamensforberedelser og prøveeksamen

Eksamensforberedelser og prøveeksamen

- Eksamensforberedelser
- Digital skoleeksamen generelt
- Spesielt for ISTx100y
- Prøveeksamen med løsning (tekst og video)
- Løsningsskisse utvalgte STACK-oppgaver
- Relevante eksamensoppgaver med videoløsning - fra søsteremnet Statistikk for sivilingeniører (TMA4240/45) - “oppgaver av vanskelighetsgrad lett- middels”

ISTA/G/T 1001/2/3
Statistikk for ingeniører

Informasjon

Informasjon om emnet

Arbeidskrav, pensum og ressurser

Referansegruppe og fagteam

Digital plenumstime

ISTA: Campus Ålesund

ISTG: Campus Gjøvik

ISTT1001/3: Campus Trondheim

ISTT1002: Campus Trondheim

Undervisning og øvinger

Fellesmodul

STACK-øvinger

Python/Jupyter

Digitalt forum

Prosjektgrupper: påmelding og kontrakt

Eksamen

Informasjon om eksamen

Formelark og tabeller

Tidligere eksamensoppgaver

Eksamensforberedelser og prøveeksamen

Kontinuasjonseksamen 17. august 2021

Kontinuasjonseksamen må holde i samme format som hovedeksamen, og dette var derfor en digital hjemmeeksamen i Inspera.

Oppgavesett: Bokmål Nynorsk

Løsningsskisse: pdf

Eksamen 16. desember 2020

Eksamen i desember 2020 hadde ulike varianter av oppgavene, for å forhindre samarbeid blant studentene.

For denne eksamenen utviklet vi flere varianter av hver oppgave:

- 9 varianter av oppgave 1
- 8 varianter av oppgave 2a
- 8 varianter av oppgave 2b
- 6 varianter av oppgave 2c
- 6 varianter av oppgave 3
- 3 varianter av oppgave 4

Vi kvalitetssikret at alle variantene var like enkle/vanskelige. For hver student ble varianter trukket tilfeldig i Inspera. Da fantes det faktisk 62 208 mulige eksamenssett!

Her er et eksempel på hvordan ett slikt eksamenssett så ut:
istx100y_eksamen_variant.pdf

Her er et utfyllende løsningsforslag til akkurat denne varianten:
istx100y_eksamen_løsningsforslag_variant.pdf

Løsningsmetodene var like for alle varianter av oppgaver, her er en oversikt for de som vil se forskjeller
istx100y_eksamen_løsning_alle_varianter.pdf

Prøveeksamener høsten 2020

Disse prøveeksamenene ble laget høsten 2020 med formål å gi eksempler på hvordan eksamensoppgavene kan se ut i Inspera (type oppgaver: hånskrevne utregninger, flervalg, tekstboks og tallsvar). Nå som dere har eksamensoppgaven fra høsten 2020 og kontinuasjonseksamen august 2021, så er disse oppgavene kanskje ikke så relevante - men vi har allikevel inkludert dem.

Hvis du har problemer med filer i blackboard så ligger også disse filene her: https://wiki.math.ntnu.no/ist100x/2020h/start

- prøveeksamen1.pdf
- løsningsforslag1.pdf
- prøveeksamen2.pdf
- løsningsforslag2.pdf

Eldre eksamensoppgaver fra emnet TALM1005

Arbeid på eksamen

- Hvis du ikke er alt for stresset: les raskt gjennom alle oppgavene så hjernen kan starte med prosessering i bakgrunnen.
- Er du stresset: bare start på oppgave 1.
- Jobb gjennom en og en oppgave, bruk kladdearket og skriv ned så mange mellomregninger som mulig, og bruk gjerne 4 desimaler hele veien. Behold gjerne mellomregninger i kalkulatoren.
- Lag skisser og tegninger så ofte du kan - da kan du kanskje se om svaret ditt er realistisk, og det blir lettere å bestemme hvilken hale du skal forholde deg til for hypotesetest.
- Skriv inn svaret i Inspera med en gang - og bruk så mange desimaler som det er spurt om!

Arbeid på eksamen

- For flervalgsoppgaver: hvis du ikke finner svaret ditt som et alternativ så er det et dårlig tegn. Veldig mange har lest korrektur på eksamen og det veldig veldig lav sannsynlighet for at det riktige svaret mangler. Regn på nytt og se om du ser tastefeil eller ulogiske overganger. Får du det ikke til så resonner deg frem til om noen alternativer er ulogiske og velg tilfeldig blant de andre.
- Marker en oppgave hvis noe gjenstår så du kan gå tilbake (eller marker de du er ferdige med - lag ditt eget system).
- Ta korte pauser og drikk/spis.

Kandidat ID
? står att

ANNA INFORMASJON:

Vend deg til ei eksamensvakt om du ønskjer å kontakte fagleg kontaktperson under eksamen. Noter spørsmålet ditt på førehand. Fagleg kontaktperson skal berre kontaktast dersom det er direkte feil eller manglar i oppgåvesettet.

Tallsvar: For kvar oppgave der du skal skrive inn eit tallsvar er det oppgitt antal desimalar du skal skrive inn. Både desimalkomma og desimalpunktum kan brukast.

Vekting av oppgåvene: er gjeve for kvar oppgåve.

Lagring: Svara dine i Inspera Assessment blir lagra automatisk kvart 15. sekund.

Varslinger: Dersom det oppstår behov for å gje beskjedar til kandidatane medan eksamen er i gang (til dømes ved klare manglar i oppgavesettet), vil dette bli gjort via varslingar i Inspera. Eit varsel vil dukke opp som ein dialogboks på skjermen i Inspera. Du kan finne att varselet ved å klikke på bjølla i øvre høgre hjørne på skjermen.

august 2022

Trekk frå eksamen: Blir du sjuk under eksamen, eller av andre grunnar ønskjer å levere blankt/trekkje deg, gå til «hamburgarmenyen» i øvre høgre hjørne og vel «Lever blankt». Dette kan ikkje angrast sjølv om prøven framleis er open. Tilkall så eksamensvakt og følg instruksane frå eksamensvakta.

Tilgang til besvarelse: Du finner besvarelsen din i Arkiv etter at sluttida for eksamen er passert.

Tabell: binomisk kumulativ s…
Tabell: Poisson kumulativ sa…
Tabell: normal kumulativ san…
Tabell: kritisk verdi i normalfo…
Tabell: kritisk verdi i t-fordelin…
Formelark