

## i Informasjon, formler og tabeller

# Eksamen

SFB12020 Anvendt statistikk og metode

6. desember 2021 kl. 9.00 - 13.00 + 15 minutter til å slutføre arbeidet med besvarelsen

Fagansvarlig: Bjørnar K. Kivedal og Irina N. Roddvik

Tillatte hjelpemidler: Godkjent kalkulator og vedlagt/utdelt formelsamling

Om oppgavesettet:

Formler og tabeller er vedlagt.

Kontroller at oppgavesettet er komplett før du begynner å besvare oppgavene.

Det er til sammen 3 -tre- oppgaver. **Alle oppgavene skal besvares og teller som angitt i parentes i oppgaveteksten ved sensurering.**

Dersom noe er uklart eller mangler i oppgavene inngår det som en del av oppgaven å ta de nødvendige forutsetninger.

Sensurfrist: **4. januar 2022**

Resultatene blir publisert i studentweb

# 1 Oppgave 1 (40%)

La oss forestille oss at du er ansatt ved et forskningsinstitutt som har fått som oppgave å gjennomføre en kvalitativ undersøkelse om **kvalitet i høyere utdanning**. **Formuler en problemstilling som du skal undersøke. Velg en eller flere institusjoner som du bruker som case for din undersøkelse og hvor du skal samle data. Begrunn kort hvorfor valg av problemstillingen.**

1. Begrunn valg av datainnsamlingsmetoder. Gjør rede for hvorfor du vil foretrekke å bruke kvalitative metoder i undersøkelsen din.
2. Hva er forskjellen mellom kvalitative og kvantitative metoder?
3. Skisser kort et kvalitativt undersøkelsesdesign du mener egner seg for å undersøke problemstillingen.
4. Du må gi klare beskrivelser av hvem du ønsker å intervju, hvordan du velger ut respondenter og hvordan du gjennomføre intervjuene.

Bruk faglige definisjoner, begrep fra kurset/pensumboken og kursmaterialet

**Skriv ditt svar her**

Format | **B** | *I* | U |  $x_2$  |  $x^2$  |  $I_x$  | | | | | | |  $\Omega$  | | |  $\Sigma$  |












Words: 0


Maks poeng: 40

## 2 Oppgave 2 (25%)

Se vedlagte PDF for oppgavetekst.

Skriv ditt svar her

Format | **B** | *I* | U |  $x_2$  |  $x^2$  |  $I_x$  |  |  |  |  |  |  |  |  |  |  |  |



Words: 0












---


Maks poeng: 25

### 3 Oppgave 3 (35%)

Se vedlagte PDF for oppgavetekst.

Skriv ditt svar her

Format ▾ | **B** *I* U  $x_2$   $x^2$  |  $I_x$  |   |    |   |   |  |  |



Words: 0

---

Maks poeng: 35

**Document 1**  
Attached



## Formler og tabeller – SFB 12020 Anvendt statistikk og metode høst 2021

Utvalgsgjennomsnittet til $X$	$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$
Utvalgsvariansen til $X$	$s_X^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$
Utvalgsstandardavviket til $X$	$s_X = \sqrt{s_X^2}$
Utvalgskovariansen mellom $X$ og $Y$	$s_{XY} = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})$
Utvalgskorrelasjonen mellom $X$ og $Y$	$r_{XY} = \frac{s_{XY}}{s_X \cdot s_Y}$
Estimerte/beregnete verdier i den enkle regresjonsmodellen	$b_2 = \frac{s_{XY}}{s_X^2}, \quad b_1 = \bar{Y} - b_2 \bar{X}$
Standardfeilen til regresjonen	$\hat{\sigma} = \sqrt{\frac{\sum \hat{u}^2}{n-k}} \quad \left( \hat{\sigma}^2 = \frac{\sum \hat{u}^2}{n-k} \right)$
Forklart kvadratsum	$ESS = \sum (\hat{Y}_i - \bar{Y})^2$
Totalkvadratsum	$TSS = \sum (Y_i - \bar{Y})^2$
Residualkvadratsum	$RSS = \sum \hat{u}_i^2 = \sum (Y_i - \hat{Y}_i)^2$
Determinasjonskoeffisienten	$R^2 = \frac{ESS}{TSS}$
Determinasjonskoeffisienten – alternativ formel ved enkel regresjon	$R^2 = r_{XY}^2$
Justert $R^2$	$\bar{R}^2 = 1 - \left[ (1 - R^2) \cdot \frac{(n-1)}{(n-k)} \right]$
Testobservator/testuttrykk til en enkel hypotesetest	$\frac{b - H_0 \text{ verdi}}{se(b)}$
Standardfeil til estimert/beregnet stigningstall. ( $R_j^2 = 0$ ved enkel regresjon)	$se(b_j) = \sqrt{var(b_j)}$ $var(b_j) = \frac{\hat{\sigma}^2}{\sum_{i=1}^n (X_{ji} - \bar{X}_j)^2} \cdot \frac{1}{1 - R_j^2}$
Et $(1 - \alpha) \cdot 100\%$ konfidensintervall for $B_i$	Øvre grense: $b_i + t_{\alpha/2}(df) \cdot se(b_i)$ Nedre grense: $b_i - t_{\alpha/2}(df) \cdot se(b_i)$
Testuttrykk til F-testen	$F = \frac{(RSS_r - RSS_{ur})/m}{RSS_{ur}/(n-k)}$
Testuttrykk til F-testen dersom $TSS_{ur} = TSS_r$	$F = \frac{(R_{ur}^2 - R_r^2)/m}{(1 - R_{ur}^2)/(n-k)}$

Frihetsgrader i hhv. teller og nevner i F-fordelingen ved multiplert hypotesetesting	$Df_1 = m$ og $Df_2 = n - k$
Regneregler eksponentialfunksjonen	$e^x \cdot e^y = e^{x+y}$ $\frac{e^x}{e^y} = e^{x-y}$ $(e^x)^y = e^{x \cdot y}$
Regneregler logaritmer	$x = e^{\ln x}, \text{ gitt at } x > 0$ $\ln 1 = 0$ $\ln e = 1$ $\ln 0 \text{ eksisterer ikke}$ $\ln(x \cdot y) = \ln x + \ln y$ $\ln\left(\frac{x}{y}\right) = \ln x - \ln y$ $\ln x^y = y \cdot \ln x$
Tilnærmet tolkning av stigningstall i log-log-sammenhenger	Dersom forklaringsvariabelen øker med 1% så endres avhengig variabel i gjennomsnitt med $b_i\%$ , cet. par.
Tilnærmet tolkning av stigningstall i log-lin-sammenhenger	Dersom forklaringsvariabelen øker med en enhet så endres avhengig variabel i gjennomsnitt med $(b_i \cdot 100)\%$ , cet. par.
Tilnærmet tolkning av stigningstall i lin-log-sammenhenger	Dersom forklaringsvariabelen øker med 1% så endres avhengig variabel i gjennomsnitt med $b_i/100$ , cet. par.
Et $(1 - \alpha) \cdot 100\%$ konfidensintervall for den faktiske verdien Y	Øvre grense: $\hat{Y} + t_{\alpha/2}(df) \cdot se(Y)$ Nedre grense: $\hat{Y} - t_{\alpha/2}(df) \cdot se(Y)$ der $se(Y) = \hat{\sigma}$ dersom anslaget er for én periode fremover i tid
Testobservator kjikvadratstest	$Q = \sum_{\text{alle celler}} \frac{(\text{observert} - \text{forventet})^2}{\text{forventet}}$
Frihetsgrader kjikvadratstest	$(r - 1)(k - 1)$
Testobservatorer	$Z = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$ $T = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}}$ $Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{X - np_0}{\sqrt{np_0(1-p_0)}}$



## Normalfordelingen

### Kumulative sannsynligheter for NEGATIVE $z$ -verdier

$z$	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,500	0,496	0,492	0,488	0,484	0,480	0,476	0,472	0,468	0,464
-0,1	0,460	0,456	0,452	0,448	0,444	0,440	0,436	0,433	0,429	0,425
-0,2	0,421	0,417	0,413	0,409	0,405	0,401	0,397	0,394	0,329	0,386
-0,3	0,382	0,378	0,375	0,371	0,367	0,363	0,359	0,356	0,352	0,348
-0,4	0,345	0,341	0,337	0,334	0,330	0,326	0,323	0,319	0,316	0,312
-0,5	0,309	0,305	0,302	0,298	0,295	0,291	0,288	0,284	0,281	0,278
-0,6	0,274	0,271	0,268	0,264	0,261	0,258	0,255	0,251	0,248	0,245
-0,7	0,242	0,239	0,236	0,233	0,230	0,227	0,224	0,221	0,218	0,215
-0,8	0,212	0,209	0,206	0,203	0,201	0,198	0,195	0,192	0,189	0,187
-0,9	0,184	0,181	0,179	0,176	0,174	0,171	0,169	0,166	0,164	0,161
-1,0	0,159	0,156	0,154	0,152	0,149	0,147	0,145	0,142	0,140	0,138
-1,1	0,136	0,134	0,131	0,129	0,127	0,125	0,123	0,121	0,119	0,117
-1,2	0,115	0,113	0,111	0,109	0,108	0,106	0,104	0,102	0,100	0,099
-1,3	0,097	0,095	0,093	0,092	0,090	0,089	0,087	0,085	0,084	0,082
-1,4	0,081	0,079	0,078	0,076	0,075	0,074	0,072	0,071	0,069	0,068
-1,5	0,067	0,066	0,064	0,063	0,062	0,061	0,059	0,058	0,057	0,056
-1,6	0,055	0,054	0,053	0,052	0,051	0,050	0,049	0,048	0,047	0,046
-1,7	0,045	0,044	0,043	0,042	0,041	0,040	0,039	0,038	0,038	0,037
-1,8	0,036	0,035	0,034	0,034	0,033	0,032	0,031	0,031	0,030	0,029
-1,9	0,029	0,028	0,027	0,027	0,026	0,026	0,025	0,024	0,024	0,023
-2,0	0,023	0,022	0,022	0,021	0,021	0,020	0,020	0,019	0,019	0,018
-2,1	0,018	0,017	0,017	0,017	0,016	0,016	0,015	0,015	0,015	0,014
-2,2	0,014	0,014	0,013	0,013	0,013	0,012	0,012	0,012	0,011	0,011
-2,3	0,011	0,010	0,010	0,010	0,010	0,009	0,009	0,009	0,009	0,008
-2,4	0,008	0,008	0,008	0,008	0,007	0,007	0,007	0,007	0,007	0,006
-2,5	0,006	0,006	0,006	0,006	0,006	0,005	0,005	0,005	0,005	0,005
-2,6	0,005	0,005	0,004	0,004	0,004	0,004	0,004	0,004	0,004	0,004
-2,7	0,004	0,003	0,003	0,003	0,003	0,003	0,003	0,003	0,003	0,003
-2,8	0,003	0,003	0,002	0,002	0,002	0,002	0,002	0,002	0,002	0,002
-2,9	0,002	0,002	0,002	0,002	0,002	0,002	0,002	0,002	0,001	0,001
-3,0	0,001	0,001	0,001	0,001	0,001	0,001	0,001	0,001	0,001	0,001
-3,1	0,001	0,001	0,001	0,001	0,001	0,001	0,001	0,001	0,001	0,001
-3,2	0,001	0,001	0,001	0,001	0,001	0,001	0,001	0,001	0,001	0,001
-3,3	0,001	0,001	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000
-3,4	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000

Verdier generert i R versjon 2.13.2 med qnorm-funksjonen



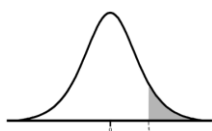
Normalfordelingen  
Kumulative sannsynligheter for POSITIVE  
 $z$ -verdier



$z$	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,500	0,504	0,508	0,512	0,516	0,520	0,524	0,528	0,532	0,536
0,1	0,540	0,544	0,548	0,552	0,556	0,560	0,564	0,568	0,571	0,575
0,2	0,579	0,583	0,587	0,591	0,595	0,599	0,603	0,606	0,610	0,614
0,3	0,618	0,622	0,626	0,629	0,633	0,637	0,641	0,644	0,648	0,652
0,4	0,655	0,659	0,663	0,666	0,670	0,674	0,677	0,681	0,684	0,688
0,5	0,692	0,695	0,699	0,702	0,705	0,709	0,712	0,716	0,719	0,722
0,6	0,726	0,729	0,732	0,736	0,739	0,742	0,745	0,749	0,752	0,755
0,7	0,758	0,761	0,764	0,767	0,770	0,773	0,776	0,779	0,782	0,785
0,8	0,788	0,791	0,794	0,797	0,780	0,802	0,805	0,808	0,811	0,813
0,9	0,816	0,819	0,821	0,824	0,826	0,829	0,832	0,834	0,837	0,839
1,0	0,841	0,844	0,846	0,849	0,851	0,853	0,855	0,858	0,860	0,862
1,1	0,864	0,867	0,869	0,871	0,873	0,875	0,877	0,879	0,881	0,883
1,2	0,885	0,887	0,889	0,891	0,893	0,894	0,896	0,898	0,900	0,902
1,3	0,903	0,905	0,907	0,908	0,910	0,912	0,913	0,915	0,916	0,918
1,4	0,919	0,921	0,922	0,924	0,925	0,927	0,928	0,929	0,931	0,932
1,5	0,933	0,935	0,936	0,937	0,938	0,939	0,941	0,942	0,943	0,944
1,6	0,945	0,946	0,947	0,948	0,950	0,951	0,952	0,953	0,954	0,955
1,7	0,955	0,956	0,957	0,958	0,959	0,960	0,961	0,962	0,963	0,963
1,8	0,964	0,965	0,966	0,966	0,967	0,968	0,969	0,969	0,970	0,971
1,9	0,971	0,972	0,973	0,973	0,974	0,974	0,975	0,976	0,976	0,977
2,0	0,977	0,978	0,978	0,979	0,979	0,980	0,980	0,981	0,981	0,982
2,1	0,982	0,983	0,983	0,983	0,984	0,984	0,985	0,985	0,985	0,986
2,2	0,986	0,986	0,987	0,987	0,988	0,988	0,988	0,988	0,989	0,989
2,3	0,989	0,990	0,990	0,990	0,990	0,991	0,991	0,991	0,991	0,992
2,4	0,992	0,992	0,992	0,993	0,993	0,993	0,993	0,993	0,993	0,994
2,5	0,994	0,994	0,994	0,994	0,995	0,995	0,995	0,995	0,995	0,995
2,6	0,995	0,996	0,996	0,996	0,996	0,996	0,996	0,996	0,996	0,996
2,7	0,997	0,997	0,997	0,997	0,997	0,997	0,997	0,997	0,997	0,997
2,8	0,997	0,998	0,998	0,998	0,998	0,998	0,998	0,998	0,998	0,998
2,9	0,998	0,998	0,998	0,998	0,998	0,998	0,999	0,999	0,999	0,999
3,0	0,999	0,999	0,999	0,999	0,999	0,999	0,999	0,999	0,999	0,999
3,1	0,999	0,999	0,999	0,999	0,999	0,999	0,999	0,999	0,999	0,999
3,2	0,999	0,999	0,999	0,999	0,999	0,999	0,999	1,000	1,000	1,000
3,3	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000
3,4	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000

Verdier generert i R versjon 2.13.2 med qnorm-funksjonen

# t-fordelingen: Kritiske verdier



Frihetsgrader (df)	Halesannsynligheter:											
	0.30	0.25	0.20	0.15	0.10	0.05	0.025	0.02	0.01	0.005	0.0025	0.001
1	0.727	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3
2	0.617	0.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33
3	0.584	0.765	0.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21
4	0.569	0.741	0.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173
5	0.559	0.727	0.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893
6	0.553	0.718	0.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208
7	0.549	0.711	0.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785
8	0.546	0.706	0.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501
9	0.543	0.703	0.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297
10	0.542	0.700	0.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144
11	0.540	0.697	0.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025
12	0.539	0.695	0.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930
13	0.538	0.694	0.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852
14	0.537	0.692	0.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787
15	0.536	0.691	0.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733
16	0.535	0.690	0.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686
17	0.534	0.689	0.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646
18	0.534	0.688	0.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.610
19	0.533	0.688	0.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579
20	0.533	0.687	0.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552
21	0.532	0.686	0.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527
22	0.532	0.686	0.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505
23	0.532	0.685	0.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485
24	0.531	0.685	0.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467
25	0.531	0.684	0.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450
26	0.531	0.684	0.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435
27	0.531	0.684	0.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421
28	0.530	0.683	0.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408
29	0.530	0.683	0.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396
30	0.530	0.683	0.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385
31	0.530	0.682	0.853	1.054	1.309	1.696	2.040	2.144	2.453	2.744	3.022	3.375
32	0.530	0.682	0.853	1.054	1.309	1.694	2.037	2.141	2.449	2.738	3.015	3.365
33	0.530	0.682	0.853	1.053	1.308	1.692	2.035	2.138	2.445	2.733	3.008	3.356
34	0.529	0.682	0.852	1.052	1.307	1.691	2.032	2.136	2.441	2.728	3.002	3.348
35	0.529	0.682	0.852	1.052	1.306	1.690	2.030	2.133	2.438	2.724	2.996	3.340
36	0.529	0.681	0.852	1.052	1.306	1.688	2.028	2.131	2.434	2.719	2.990	3.333
37	0.529	0.681	0.851	1.051	1.305	1.687	2.026	2.129	2.431	2.715	2.985	3.326
38	0.529	0.681	0.851	1.051	1.304	1.686	2.024	2.127	2.429	2.712	2.980	3.319
39	0.529	0.681	0.851	1.050	1.304	1.685	2.023	2.125	2.426	2.708	2.976	3.313
40	0.529	0.681	0.851	1.050	1.303	1.684	2.021	2.123	2.423	2.704	2.971	3.307
41	0.529	0.681	0.850	1.050	1.303	1.683	2.020	2.121	2.421	2.701	2.967	3.301
42	0.528	0.680	0.850	1.049	1.302	1.682	2.018	2.120	2.418	2.698	2.963	3.296
43	0.528	0.680	0.850	1.049	1.302	1.681	2.017	2.118	2.416	2.695	2.959	3.291
44	0.528	0.680	0.850	1.049	1.301	1.680	2.015	2.116	2.414	2.692	2.956	3.286
45	0.528	0.680	0.850	1.049	1.301	1.679	2.014	2.115	2.412	2.690	2.952	3.281
46	0.528	0.680	0.850	1.048	1.300	1.679	2.013	2.114	2.410	2.687	2.949	3.277
47	0.528	0.680	0.849	1.048	1.300	1.678	2.012	2.112	2.408	2.685	2.946	3.273
48	0.528	0.680	0.849	1.048	1.299	1.677	2.011	2.111	2.407	2.682	2.943	3.269
49	0.528	0.680	0.849	1.048	1.299	1.677	2.010	2.110	2.405	2.680	2.940	3.265
50	0.528	0.679	0.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678	2.937	3.261
51	0.528	0.679	0.849	1.047	1.298	1.675	2.008	2.108	2.402	2.676	2.934	3.258
52	0.528	0.679	0.849	1.047	1.298	1.675	2.007	2.107	2.400	2.674	2.932	3.255
53	0.528	0.679	0.848	1.047	1.298	1.674	2.006	2.106	2.399	2.672	2.929	3.251
54	0.528	0.679	0.848	1.046	1.297	1.674	2.005	2.105	2.397	2.670	2.927	3.248
55	0.527	0.679	0.848	1.046	1.297	1.673	2.004	2.104	2.396	2.668	2.925	3.245
60	0.527	0.679	0.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.915	3.232
70	0.527	0.678	0.847	1.044	1.294	1.667	1.994	2.093	2.381	2.648	2.899	3.211
80	0.526	0.678	0.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639	2.887	3.195
90	0.526	0.677	0.846	1.042	1.291	1.662	1.987	2.084	2.368	2.632	2.878	3.183
100	0.526	0.677	0.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.174
1000	0.525	0.675	0.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813	3.098
∞	0.524	0.674	0.842	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.090
	40%	50%	60%	70%	80%	90%	95%	96%	98%	99%	99.5%	99.8%

Generert i R versjon 2.13.2 med qt funksjonen.









## Kritiske verdier kjikvadratfordelingen

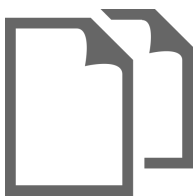


Frihets- grader	Signifikansnivå:			
	10%	5%	1%	0.1%
1	2.7055	3.8415	6.6349	10.8276
2	4.6052	5.9915	9.2103	13.8155
3	6.2514	7.8147	11.3449	16.2662
4	7.7794	9.4877	13.2767	18.4668
5	9.2364	11.0705	15.0863	20.5150
6	10.6446	12.5916	16.8119	22.4577
7	12.0170	14.0671	18.4753	24.3219
8	13.3616	15.5073	20.0902	26.1245
9	14.6837	16.9190	21.6660	27.8772
10	15.9872	18.3070	23.2093	29.5883
11	17.2750	19.6751	24.7250	31.2641
12	18.5493	21.0261	26.2170	32.9095
13	19.8119	22.3620	27.6882	34.5282
14	21.0641	23.6848	29.1412	36.1233
15	22.3071	24.9958	30.5779	37.6973
16	23.5418	26.2962	31.9999	39.2524
17	24.7690	27.5871	33.4087	40.7902
18	25.9894	28.8693	34.8053	42.3124
19	27.2036	30.1435	36.1909	43.8202
20	28.4120	31.4104	37.5662	45.3147
21	29.6151	32.6706	38.9322	46.7970
22	30.8133	33.9244	40.2894	48.2679
23	32.0069	35.1725	41.6384	49.7282
24	33.1962	36.4150	42.9798	51.1786
25	34.3816	37.6525	44.3141	52.6197
26	35.5632	38.8851	45.6417	54.0520
27	36.7412	40.1133	46.9629	55.4760
28	37.9159	41.3371	48.2782	56.8923
29	39.0875	42.5570	49.5879	58.3012
30	40.2560	43.7730	50.8922	59.7031
31	41.4217	44.9853	52.1914	61.0983
32	42.5847	46.1943	53.4858	62.4872
33	43.7452	47.3999	54.7755	63.8701
34	44.9032	48.6024	56.0609	65.2472
35	46.0588	49.8018	57.3421	66.6188
40	51.8051	55.7585	63.6907	73.4020
45	57.5053	61.6562	69.9568	80.0767
50	63.1671	67.5048	76.1539	86.6608
55	68.7962	73.3115	82.2921	93.1675
60	74.3970	79.0819	88.3794	99.6072
65	79.9730	84.8206	94.4221	105.9881
70	85.5270	90.5312	100.4252	112.3169
80	96.5782	101.8795	112.3288	124.8392
90	107.5650	113.1453	124.1163	137.2084
100	118.4980	124.3421	135.8067	149.4493
120	140.2326	146.5674	158.9502	173.6174
150	172.5812	179.5806	193.2077	209.2646
300	331.7885	341.3951	359.9064	381.4252
500	540.9303	553.1268	576.4928	603.4460
1000	1057.7239	1074.6794	1106.9690	1143.9171

Verdier generert i R versjon 2.13.2 med qchisq funksjonen.

## Question 2

Attached



## Oppgave 2 (25%)

Korte og presise svar premieres.

For å analysere sammenhengen mellom to variabler  $X$  og  $Y$ , så har vi beregnet en regresjonsmodell med utgangspunkt i et datasett bestående av  $n=25$  observasjoner. Den beregnede modellen er

$$\hat{Y} = \underset{(12,028)}{b_1} + \underset{(0,562)}{b_2} X$$

Der tallene i parentes er standardfeilene til beregningene. I tillegg har vi

$$\bar{Y} = 41,399 \quad \bar{X} = 21,144 \quad \sum_{i=1}^{25} (X - \bar{X})(Y - \bar{Y}) = 367,18$$

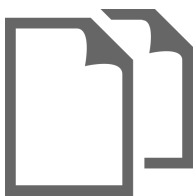
$$\sum_{i=1}^{25} (Y - \bar{Y})^2 = 2390,67 \quad \sum_{i=1}^{25} (X - \bar{X})^2 = 256,24$$

- Hva er det beregnede stigningstallet  $b_2$ ?
- Hva er det beregnede konstantleddet  $b_1$ ?
- Beregn  $R^2$
- Hva betyr verdien du beregnet i c)?
- Du skal nå gjennomføre følgende test:  $H_0: B_2 = 0$   $H_A: B_2 \neq 0$ . Bruk et 10% signifikansnivå.
  - Hva er testverdien?
  - Hva er kritisk(e) verdi(er)?
  - Hva er konklusjonen til testen?



### Question 3

Attached



### Oppgave 3 (35%)

Vi har et datasett med 3796 personer i USA. For disse personene har vi mål på variablene  $yrсед_i$  (antall års utdanning inkludert grunnskole),  $distmil_i$  (avstand i mil til nærmeste høyskole), samt  $parentscoll_i$  som er en dummyvariabel som =1 dersom minst en av foreldrene har høyere utdanning.

Nedenfor følger et datasammendrag for alle 3796 observasjoner av disse variablene.

Summary Statistics, using the observations 1 - 3796					
Variable	Mean	Median	S.D.	Min	Max
parentscoll	0,251	0,000	0,434	0,000	1,00
yrсед	13,8	13,0	1,81	12,0	18,0
distmil	2,775	1,609	3,433	0,000	25,74

- a) Du skal nå teste hvorvidt personer i gjennomsnitt har mer enn 13 års utdanning. Bruk et 1% signifikansnivå.
- i) Hvilken av disse er da en passende hypotesetest?
    - 1)  $H_0: \mu = 13 \quad H_A: \mu < 13$
    - 2)  $H_0: \mu = 13 \quad H_A: \mu > 13$
    - 3)  $H_0: \mu = 13 \quad H_A: \mu \neq 13$
  - ii) Hva er testverdi?
  - iii) Hva er kritisk(e) verdi(er)?
  - iv) Hva er konklusjonen på testen?

For å estimere hvorvidt avstand til nærmeste høyskole og foreldrenes utdanning har en effekt på hvor lang utdanning personer har, så estimerer vi modellen under

(«Modell 1»:  $yrсед_i = B_1 + B_2 distmil_i + B_3 parentscoll_i + u_i$  der  $u_i$  er et feilledd)

Model 1: OLS, using observations 1-3796  
Dependent variable: yrсед

	Coefficient	Std. Error	t-ratio	p-value	
const	13,5856	0,0406787	334,0	<0,0001	***
distmil	-0,0268006	0,00820521	-3,266	0,0011	***
parentscoll	1,26574	0,0649366	19,49	<0,0001	***
Mean dependent var	13,82929	S.D. dependent var		1,813969	
Sum squared resid	11265,88	S.E. of regression		1,723420	
R-squared	0,097819	Adjusted R-squared		0,097343	
F(2, 3793)	205,6270	P-value(F)		1,64e-85	
Log-likelihood	-7450,994	Akaike criterion		14907,99	
Schwarz criterion	14926,71	Hannan-Quinn		14914,64	

- b) Gi en tolkning av de beregnede koeffisientene i Modell 1.

I tillegg beregner vi modellen «Modell 2»  $yr_{sed}_i = B_1 + B_2 distmil_i + B_3 parentscoll_i + B_4 parentscoll_i \cdot distmil_i + u_i$  der  $u_i$  er et feilledd. I utskriften under er  $pcolldistmil_i = parentscoll_i \cdot distmil_i$

Model 2: OLS, using observations 1-3796  
Dependent variable: yrsed

	<i>Coefficient</i>	<i>Std. Error</i>	<i>t-ratio</i>	<i>p-value</i>	
const	13,5900	0,0419545	323,9	<0,0001	***
distmil	-0,0282737	0,00888499	-3,182	0,0015	***
parentscoll	1,24353	0,0827905	15,02	<0,0001	***
pcolldistmil	0,0100223	0,0231754	0,4325	0,6654	
Mean dependent var	13,82929	S.D. dependent var		1,813969	
Sum squared resid	11265,33	S.E. of regression		1,723605	
R-squared	0,097863	Adjusted R-squared		0,097149	
F(3, 3792)	137,1176	P-value(F)		2,42e-84	
Log-likelihood	-7450,901	Akaike criterion		14909,80	
Schwarz criterion	14934,77	Hannan-Quinn		14918,68	

- Gi en tolkning av den beregnede koeffisienten  $b_4$  i Modell 2.
- Beregn antall år utdanning ( $\widehat{yr_{sed}}$ ) for person nr 116 i datasettet ved hjelp av Modell 2. Foreldrene til denne personen har ikke høyere utdanning, og avstanden til nærmeste høyskole er 8,045 mil.
- Hva blir anslagsfeilen for observasjon nr 116 dersom faktisk verdi er  $yr_{sed}_{116} = 12$  år?