

Att tänka på (obligatorisk läsning)

- A. Redovisa Dina lösningar i en form som gör det lätt att följa Din tankegång. (Rättaren förutsätter att det dunkelt skrivna är dunkelt tänkt.). Motivera alla väsentliga steg i lösningen. Ange alla antaganden Du gör och alla förutsättningar Du utnyttjar
- B. Vid konfidensintervall måste Du dessutom ange vad intervallet avser att täcka samt teckna intervallet i symbolform innan de numeriska uppgifterna sätts in. Verbal slutsats av det framräknade resultatet krävs för full poäng.
- C. Vid signifikansanalys måste Du utöver vad som sagts ovan ange H_0 , H_1 , signifikansnivå (alt. P-värde), testfunktion (inkl. ev. antal frihetsgrader), resultat och verbal slutsats.
- D. Samtliga beräknade mått skall följas av en verbal slutsats för full poäng.

UPPGIFT 1

Ett stort företag har ett traineeprogram där 70 procent av traineerna är kvinnor. 90 procent av kvinnorna och 80 procent av männen är högskoleutbildade.

- a) Anta att en trainee väljs ut slumpmässigt. Hur stor är sannolikheten att denna är en kvinna som saknar högskoleutbildning? (3 p)
- b) För en viss specialutbildning vill man plocka ut en högskoleutbildad trainee. Hur stor är sannolikheten att man automatiskt får en högskoleutbildad trainee om man slumpmässigt väljer en ur högen? (3 p)
- c) Är kön och utbildningsbakgrund sinsemellan oberoende? (4 p)

UPPGIFT 2

Jörgen och Ulla ska göra en tentamensskrivning. De bestämmer sig för att skrivningen ska omfatta 6 uppgifter och definierar ungefärligt tema för respektive uppgift. Sedan låter de slumpen avgöra vem som ska göra frågorna. Låt X vara antalet uppgifter som Jörgen ska skriva.

- a) Vilken fördelning har X ? (2 p)
- b) Det visade sig att Jörgen fick skriva 2 av uppgifterna. Hur stor var sannolikheten att detta skulle inträffa? (3 p)
- c) Anta att de på samma sätt fördelar sammanlagt 300 arbetstimmar för en statistikkurs. Hur stor är då sannolikheten att Jörgen får göra högst 100 av timmarna? (5 p)

UPPGIFT 3

När ett stort hotell ska köpa in glödlampor finns det fyra tillgängliga märken. Inköpschefen begär in 100 lampor från varje tillverkare och testar dessa. Antal acceptabla respektive oacceptabla lampor från respektive leverantör redovisas i nedanstående tabell. Tyder resultatet på att det finns skillnader mellan de olika märkena?

	Tillverkare			
	A	B	C	D
Acceptabla	88	92	95	89
Oacceptabla	12	8	5	11
Totalt	100	100	100	100

UPPGIFT 4

En dricka-automat är programmerad för att i genomsnitt fylla ett glas med 20 cl dricka. Standardavvikelsen är 1 cl. Mängden dricka per glas är ungefär normalfördelad. För att kontrollera att automaten uppfyller kravet på fyllandet tar man då och då stickprov om 25 glas och mäter den genomsnittliga mängden i dessa. Man vill definiera gränserna för vad som är acceptabelt genom att programmera systemet så att det ger signal för de 5% lägsta och de 5% högsta medelvärdena.

- Ange gränserna för vad som kan anses som acceptabla medelvärden. (5 p)
- Anta att medelvärdet ökar till 21 cl. Hur stor är sannolikheten att detta upptäcks när man tar ett stickprov för att kontrollera automaten. (5 p)

UPPGIFT 5

I en studie har man försökt kartlägga sambandet mellan kön, ålder och systolisk blodtryck.

I avsnittet nedan beskrivs vad det systoliska blodtrycket är för något.

Blodtryck är det tryck som driver runt blodet i blodkärlen. Trycket mäts i mm kvicksilver (mmHg) med hjälp av en blodtrycksmanschett. Då blodet pumpas ut från hjärtat, uppstår en pulsvåg. Storleken på denna kan indirekt mätas med hjälp av manschetten. Trycket vid denna pulsvåg brukar kallas Systoliskt blodtryck ("övre trycket") och trycket, då pulsvågen gått över, brukar kallas *Diastoliskt blodtryck* ("undre trycket").

För att få kalla trycket för "högt blodtryck" krävs att blodtrycket överstiger 140/90 mmHg vid minst 3 tillfällen under 3 månader.

Högt normalt blodtryck: 130-139/85-89 mmHg.
 Mild förhöjning av blodtrycket: 140-159/90-99 mmHg.
 Måttlig förhöjning av blodtrycket: 160-179/100-109 mmHg.
 Svår förhöjning av blodtrycket: över 180/över 110 mmHg.

Det högsta värdet gäller, t.ex. är 159/105 en måttlig förhöjning av blodtrycket.

Dessa gränser används oberoende av individens ålder.

Det insamlade materialet som baseras på 30 individer kan du se nedan:

Individ	SBP	Age	Sex	Individ	SBP	Age	Sex
1	144	39	1	16	130	48	0
2	220	47	1	17	135	45	0
3	138	45	1	18	114	17	0
4	145	47	1	19	116	20	0
5	162	65	1	20	124	19	1
6	142	46	0	21	136	36	0
7	170	67	0	22	142	50	0
8	124	42	0	23	120	39	0
9	158	67	1	24	120	21	0
10	154	56	0	25	160	44	1
11	162	64	0	26	158	53	1
12	150	56	1	27	144	63	1
13	160	59	1	28	130	29	0
14	110	34	0	29	125	25	0
15	128	42	0	30	175	69	1

SBP är det systoliska blodtrycket angivet i mmHg

Age är individens ålder angivet i år

Sex är individens kön, där 0 är kvinna och 1 är man

Följande information beräknas med hjälp av Minitab:

Descriptive Statistics: SBP; Age; Sex

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1
SBP	30	0	143,20	4,16	22,80	110,00	124,75
Age	30	0	45,13	2,79	15,29	17,00	35,50
Sex	30	0	0,4333	0,0920	0,5040	0,000000000	0,000000000

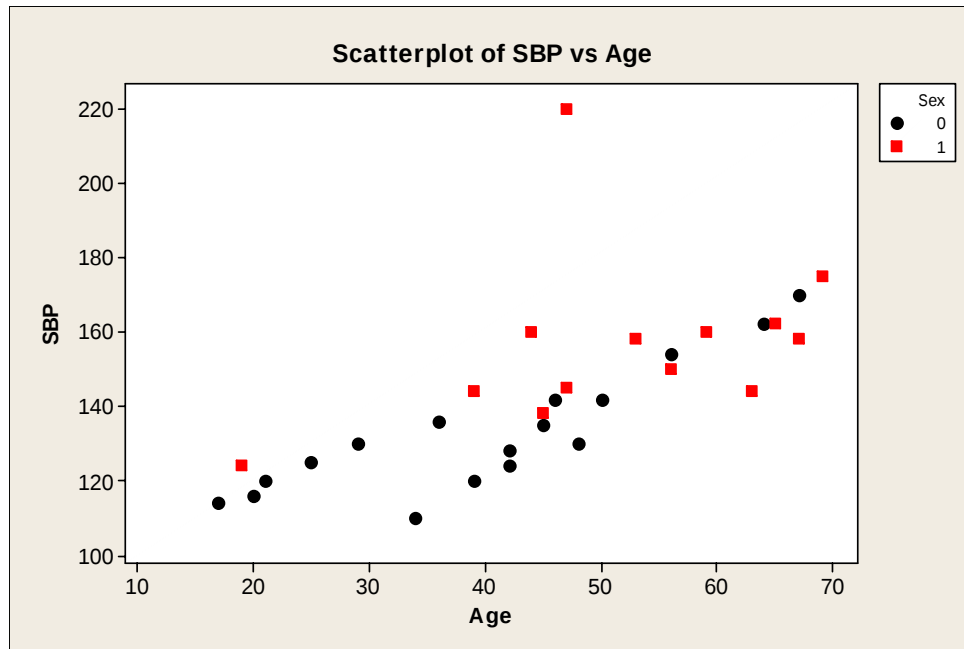
Variable	Median	Q3	Maximum
SBP	142,00	158,50	220,00
Age	45,50	56,75	69,00
Sex	0,000000000	1,0000	1,0000

- Ange för varje variabel typ och skalnivå! (2p)
- Ange vilka deskriptiva mått som är aktuella för respektive variabel! (3p)
- Vilka diagram är lämpliga för att beskriva dessa variabler? (2p)
- Beräkna lämpligt centralmått för variabeln SBP för män respektive kvinnor! (3p)

UPPGIFT 6

Uppgiften baseras på samma material som i föregående uppgift.

Ett spridningsdiagram tas fram i syfte att visualisera ett eventuellt samband mellan variablerna.



- a) Verkar det finnas ett samband mellan variablerna om man studerar spridningsdiagrammet? (2p)

För att studera sambandet mellan variablerna genomför man linjär regressionsanalys. Till att börja med så beräknar man korrelationskoefficienter mellan variablerna.

Correlations: SBP; Age; Sex

	SBP	Age
Age	0,679 0,000	

Sex	0,529 0,003	0,386 0,035
-----	----------------	----------------

Cell Contents: Pearson correlation
P-Value

- b) Vad säger dessa koefficienter? (2p)

En enkel linjär regressionsmodell med ålder som förklarande variabel ger följande resultat:

Regression Analysis: SBP versus Age

The regression equation is
 $SBP = 97,5 + 1,01 \text{ Age}$

Predictor	Coef	SE Coef	T	P
Constant	97,536	9,842	9,91	0,000
Age	1,0118	0,2069	4,89	0,000

S = 17,0387 R-Sq = 46,1% R-Sq(adj) = 44,1%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	6943,9	6943,9	23,92	0,000
Residual Error	28	8128,9	290,3		
Total	29	15072,8			

Unusual Observations

Obs	Age	SBP	Fit	SE Fit	Residual	St Resid
2	47,0	220,00	145,09	3,13	74,91	4,47R

R denotes an observation with a large standardized residual.

c) Vad säger den skattade modellen? Tolka koefficienterna!

(3p)

För att se hur kön påverkar sambandet lägger man till denna variabel till modellen.

Regression Analysis: SBP versus Age; Sex

The regression equation is

$$SBP = 99,5 + 0,831 \text{ Age} + 14,2 \text{ Sex}$$

Predictor	Coef	SE Coef	T	P
Constant	99,535	9,252	10,76	0,000
Age	0,8310	0,2098	3,96	0,000
Sex	14,215	6,367	2,23	0,034

S = 15,9422 R-Sq = 54,5% R-Sq(adj) = 51,1%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	2	8210,7	4105,3	16,15	0,000
Residual Error	27	6862,1	254,2		
Total	29	15072,8			

Source	DF	Seq SS
Age	1	6943,9
Sex	1	1266,8

- d) Hur påverkas modellen när man för in kön som en förklarande variabel? Tolka koefficienterna! Hur påverkar kön det systoliska blodtrycket? (3p)