

Tentamen (TEN1) i **Statistik och sannolikhetslära** (TNIU66) 24 augusti 2024, kl. 8.00 – 12.00.

---

### Kursens förväntade läranderesultat enligt kursplanen

Efter genomförd kurs ska du kunna:

1. analysera och visualisera fördelningen hos en datamängd.
  2. beräkna sannolikheter för vissa händelser med hjälp av teoretiska begrepp som ingår i kursinnehållet.
  3. beräkna punktskattningar och konfidensintervall.
  4. genomföra hypotesprövning
  5. genomföra enkel, linjär regressionsanalys.
  6. använda datorstöd för beräkningar där det är relevant.
- 

### Tillåtna hjälpmedel:

- Valfri bok inom statistik och sannolikhetslära<sup>1</sup>
- Miniräknare av valfritt slag (utan wifi-uppkoppling)

Det får finnas anteckningar och markeringar i boken, inklusive ”pagemarkeringar” (några centimeter stora), men inga lösblad eller inklistrade sidor.

Frågor besvaras av Michael Hörnquist som besöker skrivsalen cirka kl. 9.00 och kl. 10.30. Svar och kortfattade lösningsförslag finns på Studieinfo senast kl. 15 på tentamensdagen. Skrivningsresultat meddelas senast femton arbetsdagar efter tentamenstillfället.

Varje uppgift ger 0 – 6 poäng. Ej behandlad uppgift ges en (1) poäng, för att markera betydelsen av att veta att man inte vet. Eventuell erhållen bonus från UPG1 påförs vid rättningen och ingår i den totala poängsumman. För betyget  $n$  krävs minst  $6n - 1$  poäng, varav minst 12 poäng från tentamen. Inget övrigt krav på fördelningen av poängen föreligger.

Svaren som lämnas in ska anges på bifogad svarsblankett, och poängsättningen kommer att utgå från att det verkligen står ett svar på den. De resonemang och den kalkyl som lett fram till givna svar bifogas svarsblanketten.

**Lycka till!**

---

<sup>1</sup>Kurslitteraturen ”Tillämpad statistik – en grundkurs”, Wahlin, Sanoma förlag, torde vara vanligast.

1. Man vill bestämma medellönen för kvinnor respektive män vid ett företag med fyra avdelningar. Först samlar man in data per avdelning, medellönerna angivna i enhet "tusen kronor", och erhåller:

	Antal kvinnor	Medellön	Antal män	Medellön
Avdelning 1	22	26	7	20
Avdelning 2	34	28	14	26
Avdelning 3	14	38	32	32
Avdelning 4	9	46	26	44

- (a) På vilka avdelningar har kvinnor högst medellön?  
 (b) Vad är medellönen för samtliga kvinnor?  
 (c) Vad är medellönen för samtliga män?
2. Låt den kontinuerliga slumpvariabeln  $X$  ha täthetsfunktionen
- $$f(x) = \begin{cases} x^3/4 & \text{för } 0 \leq x \leq 2 \\ 0 & \text{för övrigt.} \end{cases}$$
- (a) Bestäm talet  $a$  så att  $\Pr(X \leq a) = 3/4$ .  
 (b) Bestäm den betingade sannolikheten  $\Pr(X \geq 1 \mid X \leq 3/2)$ .
3. På en tentamen i statistik och sannolikhetslära förekommer en uppgift med sex stycken flervalsfrågor, där var och en skall besvaras med "ja" eller "nej". Varje rätt svar ger en poäng och varje fel svar minus en poäng, dock kan totala poängsumman inte bli mindre än noll. Tentanden Sreymean är hyfsat förberedd, och har sannolikheten 75% att svara rätt på varje enskild fråga, oberoende av de andra frågorna.
- (a) Bestäm sannolikheten att hon svarar rätt på alla sex frågorna.  
 (b) Bestäm sannolikheten att hon får noll poäng totalt på uppgiften.
4. En vanlig *månad* på en förlossningsavdelning i Sverige råkar det födas 172 barn, varav 80 flickor och 92 pojkar. Bestäm ett approximativt 95% konfidensintervall för andelen flickor bland de barn som föds (a) i Sverige, (b) i EU.  
 Antag att i Sverige föds *årligen* cirka hundratusen barn, i EU cirka fyra miljoner barn.

5. Vilket eller vilka påståenden är sanna? Vilket eller vilka är falska? Markera för varje påstående "S" om det är sant och "F" om det är falskt. Lämna blankt om du är osäker.
- Det så kallade  $p$ -värdet anger sannolikheten att alternativhypotesen är falsk.
  - Det så kallade  $p$ -värdet anger sannolikheten att alternativhypotesen är sann.
  - Det så kallade  $p$ -värdet tas fram med värden som om alternativhypotesen vore sann.
  - Ett sätt att göra ett dubbelsidigt konfidensintervall dubbelt så brett är att dubblera storleken på stickprovet.
  - Ett sätt att göra ett dubbelsidigt konfidensintervall dubbelt så brett är att dubblera konfidensnivån.
  - Ett sätt att göra ett dubbelsidigt konfidensintervall dubbelt så brett är att byta till enkelsidigt.

Endast svar krävs i denna uppgift. Varje rätt svar ger en poäng och varje fel svar minus en poäng, dock kan totala poängsumman inte bli mindre än noll. Om du lämnar blankt blir det varken plus eller minus.

6. Månadsavgiften för de bostadsrätter som fanns med till salu i lokaltidning 12 oktober 2023 var:

Lägenhet	A	B	C	D	E	F	G	H
Rum (antal)	1	1	2	2	3	3	4	4
Månadsavgift (kr)	1800	1160	2130	2650	2530	2850	3100	3780

En mäklarfirma är intresserad av hur priserna varierar med antal rum och önskar dessutom veta hur pass säkra förutsägelser de kan göra för att kunna anpassa sin provision. För att åstadkomma detta genomför man en regressionsanalys. Data ovan stoppas in i Excel och dess dataanalysverktyg för regression, med antal rum som förklaringsvariabel,  $x$ , och månadsavgift som responsvariabel,  $y$ , varvid följande utdata erhålls:

UTDATASAMMANFATTNING						
<i>Regressionsstatistik</i>						
Multipel-R	0,91450					
R-kvadrat	0,83631					
Justerad R-kvadrat	0,80902					
Standardfel	352,98					
Observationer	8					
ANOVA						
	<i>fg</i>	<i>KvS</i>	<i>Mkv</i>	<i>F</i>	<i>p-värde för F</i>	
Regression	1	3819240	3819240	30,653647	0,001464	
Residual	6	747560	124593,33			
Totalt	7	4566800				
	<i>Koefficienter</i>	<i>Standardfel</i>	<i>t-kvot</i>	<i>p-värde</i>	<i>Nedre 95%</i>	<i>Övre 95%</i>
Konstant	955	305,69	3,12	0,0205	207,01	1702,99
X-variabel 1	618	111,62	5,54	0,0015	344,87	891,13

Svara nu på följande:

- Hur mycket kostar i medeltal per månad ytterligare ett rum i lägenheten?
- Bestäm ett 95% dubbelsidigt prognosintervall för vad den förväntade avgiften för en trerumslägenhet på orten kan vara.