

---

BIOSTATISTISK GRUNDKURS, MASB11, VT-19  
ÖVNING 9 (2019-05-13) OCH INFÖR ÖVNING 10

---

Aktuella avsnitt i boken: Kap. 8 och 9.

Övningens mål: Du ska:

- kunna genomföra hypotestest för proportioner med hjälp av normalapproximation
- genomföra hypotestest i en diskret fördelning utan normalapproximation
- analysera kategoridata med  $\chi^2$ -test.

1	När man studerar hur stor andel av ett stickprov som har en viss egenskap, är binomialfördelningen en väldigt vanlig modell. Parametern $p$ i binomialfördelningen kan då tolkas som proportionen i populationen med den intressanta egenskapen. Vi kan t.ex. vilja göra konfidensintervall för $p$ . I det fall stickprovet är så stort att binomialfördelningen kan approximeras med en normalfördelning ska intervallet göras som på s. 185 i boken. Gör uppgift 5.55 i arbetsmaterialet.
2	Det är inte ovanligt att man vill jämföra proportionerna i två populationer – ”är proportionen överviktiga (hur det nu definieras) större i Skåne än i Småland?” Hur man jämför $p_1$ och $p_2$ , i stora stickprov, genom att göra ett konfidensintervall för $p_1 - p_2$ beskrivs på s. 188. Gör uppgift 5.78 i arbetsmaterialet.
3	När stickprovet är litet i en binomialfördelning kan man inte ”på vanligt sätt” göra konfidensintervall eller test beträffande $p$ . Men $P$ -värdet (prob-värdet) kan beräknas och det är ett smidigt sätt att utföra testet på. Läs gärna exempel 8.7 i boken. Gör uppgift Dig:5.3_3, Dig:5.3_4 och 5.54.
4	Att använda direktmetoden, d.v.s. beräkna $P$ -värdet är lika smidigt när man sätter upp hypoteser om $\lambda$ i en poissonfördelning. Gör uppgift Dig:5.3_5 och 5.57.
5	Gör helt klart för er vad kategoridata är!
6	Vid analys av kategoridata används ofta s.k. $\chi^2$ -test. Det finns emellertid en rad av olika test som sammanfattas under namnet $\chi^2$ -test – här stöter vi på ”test av modellanpassning”, ”homogenitetstest” och ”test av oberoende”. Även om testförfaringsättet är väldigt lika i de tre fallen så är det viktigt att ha klart för sig <b>vad</b> man testar – d.v.s. studera nollhypoteserna i de tre fallen!
7	Test av modellanpassning: Här vill man testa om data kan tänkas komma från en viss fördelning. Läs exempel 9.1. Gör uppgift Dig:5.4_1 och _2.
8	Homogenitetstest: Nu vill man jämföra om $r$ stickprov har samma fördelning. Läs exempel 9.3 och observera att nollhypotesen är att man har <b>samma</b> sannolikhet i samtliga stickprov att hamna i en grupp – men vi är inte intresserade av <b>hur</b> stor den sannolikheten är. Observera också att vi vet stickprovets storlek, d.v.s. antingen rad- eller kolumnsummorna (beroende på hur tabellen presenteras) är kända <b>innan</b> vi utför experimentet (jfr. oberoendetestet). Gör uppgift 5.81 i studiematerialet. Observera att ni löst denna uppgift tidigare fast då genom att jämföra två proportioner. Det visar sig att de två testen är ekvivalenta — se s. 216–217 i boken.
	<i>Fortsätter på nästa sida!</i>

VÄND!

<b>9</b>	Oberoendetest: Studera exempel 9.4 och observera speciellt nollhypotesen: ”de två studerade variablerna är oberoende”. Notera också att rad- och kolumnsummorna inte är kända innan man gjort undersökningen. Det enda som forskarna i exempel 9.4 vet innan de tittar på data är att man studerar 2484 personer – <b>sedan</b> fördelar sig dessa personer i olika kategorier beträffande hjärtproblem och snarkbeteende. Gör uppgift Dig:5.4_3, Dig:5.4_4 och 5.83.
	Om du vill <b>träna mer</b> på detta avsnitt eller när du <b>repeterar</b> är följande uppgifter lämpliga att titta på: <b>Dig:5.3_2, 5.56, 5.64, 5.82</b>

### Inför övning 10 (2019-05-17):

<b>A</b>	Läs igenom 11.1 och studera exemplen.
<b>B</b>	<p>I avsnitt 11.2:</p> <ul style="list-style-type: none"> <li>• 11.2.1 och 11.2.2 är viktigt — studera speciellt vilka antaganden vi gör om modellen.</li> <li>• Skumma igenom 11.2.3 men fördjupa dig inte i formlerna — vi låter R eller miniräknaren göra grovjobbet. <b>Tag reda på hur du får fram skattningarna på din miniräknare!</b></li> <li>• Läs igenom 11.2.4 och koncentrera dig på <b>varför</b> det är intressant att testa <math>H_0 : \beta_1 = 0</math>.</li> <li>• Koncentrera dig i avsnitt 11.2.5 på frågeställningarna i avsnitten ”<i>konfidensintervall ...</i>” och ”<i>prediktionsintervall ...</i>” — försök förstå skillnaden mellan de två frågeställningarna.</li> </ul>
<b>C</b>	Läs igenom avsnitt 11.3. Korrelationskoefficienten, som mäter det linjära sambandet mellan två variabler, används ofta i olika sammanhang — och missbrukas också ofta!