

TENTAMEN: Matematisk statistik och diskret matematik (MVE050/MSG810). Statistik för fysiker (MSG820).

Tid och plats: Onsdagen den 27 april 2011, kl. 08.30–12.30, Väg och vatten.

Jour: Stefan Erikshed, tel 0736281687.

Tillåtna hjälpmedel: Chalmersgodkänd räknare och Beta.

Betygsgränser: Chalmers: 3: 12 poäng, 4: 18 poäng, 5: 24 poäng. GU: G: 12 poäng, VG: 21 poäng. Maximalt antal poäng är 30.

1. (3p) Ange lämplig typ av fördelning för de stokastiska variabler som beskrivs nedan.
 - a) Antal femmor man får när man slår sex tärningar.
 - b) Massan hos en slumpvald limpa från en brödproducent.
 - c) En diskret stokastisk variabel med frekvensfunktionen $f(x) = 2^{-x}$, $x = 1, 2, 3, \dots$
 - d) Tiden mellan två partikelemissioner från ett radioaktivt preparat.
 - e) Antal anrop till en webserver under en tvåminutersperiod.
 - f) En kontinuerlig stokastisk variabel med täthetsfunktion

$$f(x) = \begin{cases} \frac{1}{20}, & 30 \leq x \leq 50 \\ 0, & \text{annars.} \end{cases}$$

2. (3p) En tvådimensionell stokastisk variabel (X, Y) har följande frekvenstabell.

		y	
		0.5	2
x	1	0.30	0.20
	2	0.13	0.06
	3	0.01	0.30

- a) Beräkna $\mathbf{E}[X]$ och $\mathbf{E}[Y]$.
 - b) Beräkna $\mathbf{Cov}(X, Y)$.
 - c) Är X och Y oberoende? Motivera.
3. (3p) En maskin tillverkar en viss typ av detaljenheter. Detaljernas klassas som korrekta, acceptabla eller defekta. Om en tillverkad detalj är korrekt är sannolikheten att nästa är korrekt 0.95, och sannolikheten att den är acceptabel är 0.05. Om en tillverkad enhet är acceptabel är sannolikheten 0.84 att även nästa är acceptabel, medan sannolikheten att nästa är defekt är lika stor som sannolikheten att den är korrekt. När en defekt detalj tillverkats stoppas maskinen automatiskt, och en manuell justering måste göras varefter nästa enhet blir korrekt. Vad är väntevärdet för antal tillverkade icke-defekta enheter mellan två manuella justeringar?
4. (4p)
 - a) Formulera Centrala gränsvärdessatsen (CGS).
 - b) Visa med hjälp av CGS att binomialfördelningen under vissa förutsättningar kan approximeras med en normalfördelning.

5. (4p) En bredbandsleverantör hävdar att väntevärdet för ping-tiden mätt i millisekunder, X , från en viss typ av dator är 9.0. Rut vill testa detta påstående genom att utföra följande hypotestest.

$$H_0 : \mu_X = 9.0$$

$$H_1 : \mu_X > 9.0$$

Från ett stickprov av storlek 120 erhålls $\sum_{i=1}^{120} x_i = 1092$ och $\sum_{i=1}^{120} x_i^2 = 10900$.

- Vilken teststatistika kan Rut använda, och vad är dess fördelning under H_0 ? Motivera.
 - Vad menas med P -värdet för ett hypotestest?
 - Beräkna P -värdet för Ruts test.
 - Verkar det rimligt att den förväntade ping-tiden är 9.0ms?
6. (3p) Hur många heltalslösningar har följande ekvation?

$$\begin{cases} y_1 + y_2 + y_3 = 28 \\ y_1 \geq 8, 1 \leq y_2 \leq 10, y_3 \geq 1 \end{cases}$$

Du kan behöva serien $\sum_{n=0}^{\infty} \binom{n+k}{k} x^n = \frac{1}{(1-x)^{k+1}}$.

7. (3p)

- Definiera den momentgenererande funktionen (mgf) till en stokastisk variabel X .
- En diskret stokastisk variabel X har mgf

$$m_X(t) = (0.7e^t + 0.3)^6.$$

Beräkna $\mathbf{E}[X]$.

8. (4p) I en opinionsundersökning inför ett riksdagsval tillfrågas ett stickprov på 3300 röstberättigade vilket parti de ska rösta på. 149 personer svarar Centern.
- Gör en punktskattning av andelen röster, p , som Centern har i hela riket.
 - Vad innebär det att en skattare är väntevärdesriktig? Är skattaren du använt för p i a) väntevärdesriktig? Motivera.
 - Beräkna ett dubbelsidigt konfidensintervall av grad 99% för p .
 - För att få sitta med i riksdagen behöver ett parti minst 4% av rösterna. Utifrån dina beräkningar, har Centern skäl att vara oroliga att inte klara detta?

9. (3p) Den globala uppvärmningen är ett aktuellt problem. En väderstation i en ort har gjort mätningar av årsmedeltemperaturen sedan 1910. Följande tabell ger temperaturökningen, y i $^{\circ}C$, x är tiden i år efter 1910.

År	x	Temp. ökn. [$^{\circ}C$]
		y
1930	20	0.02
1950	40	0.05
1970	60	0.07
1980	70	0.11
2000	90	0.15
2010	100	0.19

- Antag att uppvärmningen på orten skett ungefär linjärt under åren 1910 till 2010, och skatta den linjära regressionslinjen, $\mu_{y|x} = \beta_0 + \beta_1 x$.
- Använd din modell för att skatta temperaturökningen mellan år 1910 och 2020.
- Varför är det ytterst tveksamt att använda modellen för att skatta temperaturökningen till år 2090?

Lycka till!