Svar till tentamen i Matematisk statistik IT (TMS155), 18 december 2004

OBS: Detta är bara kortfattade, ibland ofullständiga, svar, dvs långt ifrån hur lösningarna på en tenta ska se ut!

1. a) Geometrisk b) Normalfördelning c) Poisson d) Binomial
   e) Om X är bin(10,0.2) så är $P(X \geq 7) = \sum_{i=7}^{10} \binom{10}{i} 0.2^i 0.8^{10-i} \approx 0.00086$.

2. a) A och B är ober om $P(A \cap B) = P(A)P(B)$.
   b) $P(A) = 1/6$, $P(B) = 5/36$, $P(C) = 1/6$,
   $P(A \cap B) = 1/36 \neq P(A)P(B) \rightarrow$ beroende,
   $P(B \cap C) = 0 \neq P(B)P(C) \rightarrow$ beroende,
   $P(A \cap C) = 1/36 = P(A)P(C) \rightarrow$ oberoende.

3. a) Typ I Fel: Att förkasta H_0 då H_0 är sann. Typ II fel: Att inte förkasta H_0 då H_0 är falsk.
   b) Oberoende observationer, vilken är populationen och täcker stickprovet in hela pop, är data normalfördelat, tillräckligt stort stickprov om ej normalförd, vad beror bortfall på, mm.

4. Låt X_i = höjd låda i. X_i'na ober N(65, $\sqrt{3}$). Då är $X_1 + X_2 + X_3$ $N(195, 3)$
   $P(100 + X_1 + X_2 + X_3 \leq 300) = P(X_1 + X_2 + X_3 \leq 200) = \Phi((200 - 195)/3) = \Phi(5/3) \approx$ 0.9525.

5. Tillräckligt stora stickprov för normalapprox:
   $\hat{p}_A - \hat{p}_B$ approx $N(p_A - p_B, \sqrt{p_A(1-p_A)/50 + p_B(1-p_B)/50})$.
   Testa $H_0 : p_A \leq p_B$ mot $H_1 : p_A > p_B$.
   Låt $\hat{p} = (X_A + X_B)/100$. Teststatistiska:
   $(\hat{p}_A - \hat{p}_B)/\sqrt{\hat{p}(1-\hat{p})(2/50)} = (40/50 - 33/50)/\sqrt{73/100(1 - 73/100)2/50} \approx 1.577$
   p-värde $= P(Z \geq 1.577) \approx (1 - 0.94) = 0.06$. Kan ej förkasta H_0.

6. a) Måste vänta mer än 5 minuter om jag anländer i ngt av intervallen: [7.40, 7.45], [7.50,7.55],[8.00,8.05],[8.10,8.15], dvs under 20 min av de 40 möjliga. Likformig fördelning ger P(vänta mer än 5 min) = 1/2.
   b) En buss anländer likformigt i [7.55,8.05] och en i [8.05,8.15]. Jag måste vänta mer än 5 min om den första kommer under [7.55,8.00], dvs med slh 1/2.

7. $F_Y(y) = P(Y \leq y) = P(\cap_{i=1}^n \{X_i \leq y\}) = \prod_{i=1}^n P(X_i \leq y) = (1 - (1-p)^y)^n$.
   Ober utnyttjas i 3dje likheten ovan.
   $f_Y(y) = F_Y(y) - F_Y(y-1) = (1 - (1-p)^y)^n - (1 - (1-p)^{y-1})^n$. (Se skiplistarbete.)

8. a) Om månaderna numreras $1, 2, \ldots, 13$ så $\hat{\beta}_0 = 1.57$, $\hat{\beta}_1 = 0.0141$, $\hat{\sigma}^2 = 0.000367$.
   0,...,12: 1.58+0.0141 x;
   b) Hög förklaringsgrad ($R^2 \approx 0.90$), men residualerna tyder på att en linjär modell inte är korrekt (negativa i början o slutet, positiva i mitten).

9. Se bok s 237.

10. Poissonprocess med intensitet λ kunder/min. Antal kunder i en 15-minutersperiod är då Poissonfördelat med parameter $\theta = 15\lambda$. En skattning av θ är medelvärdet $(7 + 12 + \cdots + 10)/10 = 8.9$, vilket ger $\hat{\lambda} = 8.9/15 \approx 0.5933$. Tiden tills första kunden anländer är exponentialfördelad med väntevärde $1/\lambda$ vilket kan skattas med $1/\hat{\lambda} \approx 1.69$.