Lösningar till tentamen i Matematisk statistik D (TMA290) och IT (TMS155), 29 mars 2005

1. a) $E[X] = 3 \cdot 0.5 + 2 \cdot 0.2 + 1 \cdot 0.1 = 2$, $Var(X) = (3-2)^2 0.5 + ... + (0-2)^2 .2 = 1.4$

   b) Centrala gränsvärdessatesen ger att summan är approximativt normalfördelad med väntevärde $E[\sum_{i=1}^{100} X_i] = 100E[X_1] = 300$, och varians $Var(\sum_{i=1}^{100} X_i) = 100Var(X_1) = 140$.

2. a) p-värde= P(att man får ett minst lika extremt utfall som det observerade $|H_0$ är sann). Det används för att besluta om man ska förkasta H_0 eller ej. Ju mindre p-värde desto större större skäl har man att förkasta H_0. Vanliga gränser då man förkastar är 0.01 och 0.05, men det beror på flera saker, tex vilka konsekvenser det kan få om man förkastar felaktigt.

   b) Ett $100(1-\alpha)\%$ konfidensintervall för en paramter θ är ett intervall $[L_1, L_2]$ baserat på ett stickprov sådant att $P(L_1 \leq \theta \leq L_2) = 1 - \alpha$.

   Det är alltså ett osäkerhetsintervall för en skattning. Om man tar nya stickprov 100 gånger och varje gång konstruerar tex ett 95% konfidensintervall kommer ca 95 st av de 100 intervallen att innehålla det sanna värdet på parametern θ.

3. $$P(X > y) = \int_y^\infty abx^{b-1}e^{-ax^b}dx = [-e^{-ax^b}]_y^\infty = e^{-ay^b}.$$

   $$P(X > t + 1.5|X > t) = P(X > t + 1.5)/P(X > t) = e^{-a(t+1.5)^b}/e^{-at^b}$$

   $t = 3 \Rightarrow \approx 0.962$
   $t = 9 \Rightarrow \approx 0.885$

4. $\bar{x} \approx 9.09$, $s^2 \approx 4.05$

   Konfidensintervall: $[15s^2/25, 15s^2/7.26]$ där 25 och 7.26 fås ur χ^2_{15}-tabell $\Rightarrow [2.43, 8.37]$.

5. a) $Kov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)] = E[XY - X\mu_Y - Y\mu_X + \mu_X\mu_Y] =$ $E[XY] - \mu_Y E[X] - \mu_X E[Y] + \mu_X\mu_Y = E[XY] - \mu_X\mu_Y$.

   b) Diskreta fallet: Ober betyder att $f_{X,Y}(x,y) = f_X(x)f_Y(y)$ för alla x, y, vilket ger $E[XY] = \sum_x \sum_y xyf_{X,Y}(x,y) = \sum_x xf_X(x) \sum_y yf_Y(y) = E[X]E[Y]$.

6. a) $P(A \cap B \cap C) = P(A)P(B)P(C) = 0.1^3 = 0.001$ pga ober.
   b)
   $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C) =$
   $P(A) + P(B) + P(C) - P(A)P(B) - P(A)P(C) - P(B)P(C) + P(A)P(B)P(C) = 0.271$
   c) $P(A \cup B \cup C) \leq P(A) + P(B) + P(C) = 0.3$

7. Låt X_i = viktökning fisk nr i miljö 1, Y_i = viktökning fisk nr i miljö 2.

   $\bar{x} = 59.375, \bar{y} = 45.143, s_x^2 = 160.554, s_y^2 = 54.476, n_x = 8, n_y = 7$

   $s_p^2 = ((n_x - 1)s_x^2 + (n_y - 1)s_y^2)/(n_x + n_y - 2) = 111.595$

   $H_0 : \mu_x = \mu_2 \Leftrightarrow \mu_x - \mu_y = 0$
   $H_1 : \mu_x \neq \mu_2 \Leftrightarrow \mu_x - \mu_y \neq 0$

   Teststatistika: $T = (\bar{X} - \bar{Y})/(s_p\sqrt{1/n_x + 1/n_y})$

   Observerat $T \approx 2.60$.

   p-värde= $P(|T| > 2.6|H_0$ är sann$) < 0.01$, fås ur T_{13}-tabell. Vi kan förkasta H_0 och dra slutsatsen att det finns en skillnad i viktökning.

8. Låt X = antal impulser/halvtimma $\sim Po(\lambda)$. Vid bakgrundsstrålning är $\lambda = 2744/2 = 1372$.
   $H_0 : \lambda = 1372$
   $H_1 : \lambda > 1372$

   X är approximativt normalfördelad med väntevärde och varians λ.

   p-värde= $P(X > 1503|H_0$ är sann$) = P((X - 1372)/\sqrt{1372} > (1503 - 1372)/\sqrt{1372}) \approx 1 - \Phi(3.54) \approx 0.001$. Vi kan förkasta H_0.

9. Låt X= startpunkt $\sim N(1, 2)$, Y= förflyttning $\sim N(2, 3)$, $W = X + Y$= slutpunkt. Eftersom X och Y är oberoende blir W normalfördelad med väntevärde 1+2=3 och varians 4+9=13.

   $P(W \leq 0) = P((W - 3)/\sqrt{13} \leq -3/\sqrt{13}) = \phi(-0.831) = 0.2033$.

10. Låt X=längsta biten och låt U vara likformigt fördelad på [0,L].
    Om $x < L/2$ så är $F_X(x) = P(X \leq x) = 0$ och om $x > L$ så är $F_X(x) = P(X \leq x) = 1$.
    För $L/2 < x < L$ gäller

    $F_X(x) = P(X \leq x) = 1 - P(X > x) = 1 - P(\{U < L - x\} \cup \{U > x\}) = 1 - 2(L - x)/L = 2x/L - 1$.

    Derivering ger $f_X(x) = 2/L$, för $L/2 < x < L$, och 0 för övrigt, dvs längsta biten är likformigt fördelad på $[L/2, L]$.