1)a) $X$ = antalet dragningar

Möjliga utfall

$(r_1 r_2 s_1 s_2,\ r_1 r_2 s_2 s_1,\ r_2 r_1 s_1 s_2,\ r_2 r_1 s_2 s_1,$

$r_1 s_1 r_2 s_2,\ r_1 s_2 r_2 s_1,\ r_2 s_1 r_1 s_2,\ r_2 s_2 r_1 s_1,$

$s_1 r_1 r_2 s_2,\ s_2 r_1 r_2 s_1,\ s_1 r_2 r_1 s_2,\ s_2 r_2 r_1 s_1,$

$r_1 s_1 s_2 r_2,\ \ldots,\ \ldots,\ \ldots,$

$s_1 r_1 s_2 r_2,\ \ldots,\ \ldots,\ \ldots,$

$s_1 s_2 r_1 r_2,\ \ldots,\ \ldots,\ \ldots)$

Totalt 24 utfall

$P_X(2) = 4/24 = 1/6$

$P_X(3) = 8/24 = 1/3$

$P_X(4) = 12/24 = 1/2$

b) $E[X] = \sum_{i=2}^{4} x\, P_X(x) = 2 \cdot 1/6 + 3 \cdot 1/3 + 4 \cdot 1/2 =$

$= 1/3 + 1 + 2 = 10/3 = 3{,}33$

c) $Var[X] = E[X^2] - (E[X])^2$

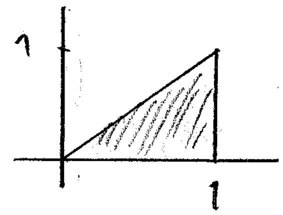
$E[X^2] = \sum_{i=2}^{4} x^2 P_X(x) = 2^2\, 1/6 + 3^2\, 1/3 + 4^2\, 1/2 =$

$= 2/3 + 3 + 8 = 35/3$

$Var[X] = 35/3 - (10/3)^2 = 5/9 \qquad \sigma = \sqrt{Var[X]} = 0{,}745$

2a) $f_X(x) = 1 \quad 0 \le x \le 1 \quad \text{Uni}(0,1)$

$f_{Y|X}(y|x) = 1/x \quad 0 \le y \le x \quad \text{Uni}(0,x)$



$f_{X,Y}(x,y) = f_{Y|X}(y|x)\, f_X(x) = 1/x \quad 0 \le y \le x \le 1$

b) $E[Y] = E[E[Y|X]] = E\left[\frac{X}{2}\right] = \frac{1}{2}E[X] = \frac{1}{2}\cdot\frac{1}{2} = 1/4$

c) $\text{Kov}(X,Y) = E[XY] - E[X]E[Y]$

$E[XY] = \int_0^1 \int_0^x xy \frac{1}{x}\, dy\, dx = \int_0^1 \int_0^x y\, dy\, dx =$

$= \int_0^1 \left[\frac{1}{2}y^2\right]_0^x dx = \frac{1}{2}\int_0^1 x^2 dx = \frac{1}{2}\cdot\frac{1}{3}\left[x^3\right]_0^1 = 1/6$

$\text{Kov}(X,Y) = 1/6 - 1/2 \cdot 1/4 = 1/6 - 1/8 = 1/24$

3 a) $Y = \sum_{i=1}^{30} X_i$

$X_i \sim Po(1/3)$

$E[X_i] = 3$

$Var[X_i] = 9$

$Y \sim N(\mu, \sigma^2)$ opproximativt enligt CLT ($n \geq 30$)

$\mu = E[Y] = 3 \cdot 30 = 90$

$\sigma^2 = Var[Y] = 9 \cdot 30 = 270$

$Y \sim N(90, 270)$

b) $P(|Y-90| \geq 10) = 1 - P(|Y-90| \leq 10)$

$P(|Y-90| \leq 10) = P(-10 \leq Y - 90 \leq 10) =$

$= P(80 \leq Y \leq 100) = P\left(\frac{80-90}{\sqrt{270}} \leq \frac{Y-90}{\sqrt{270}} \leq \frac{100-90}{\sqrt{270}}\right) =$

$= P(-0{,}609 \leq \underset{\substack{\uparrow \\ N(0,1)}}{Z} \leq 0{,}609) =$

$= \Phi(0{,}609) - \Phi(-0{,}609) =$

$= 2\Phi(0{,}609) - 1 = 0{,}46$

$P(|Y-90| \leq 10) = 1 - 0{,}46 = 0{,}54$

4. $X$ = antal fiskar med skador i art A $\quad X \sim Bin(n_A, p_A)$
$Y$ = -"- B $\quad Y \sim Bin(n_B, p_B)$

$n_A = 200$
$n_B = 300$ $\quad$ observationer $x = 30$, $y = 70$

a) $E(X) = n_A p_A \Rightarrow \hat{p}_A = \frac{x}{n_A} = \frac{30}{200} = 0.15$
$E(Y) = n_B p_B \Rightarrow \hat{p}_B = \frac{y}{n_B} = \frac{70}{300} \approx 0.23$

$E(p_A^*) = E(\frac{X}{n_A}) = p_A$ vvr. $\quad$ pss för B

b) $p_A^* = \frac{X}{n_A}$ appr $N(p_A, \frac{p_A(1-p_A)}{n_A})$ $\quad$ pss för $p_B^*$

Sökt: konfidensintervall för $p_A - p_B$:

$$I_{p_A - p_B} = \left(\hat{p}_A - \hat{p}_B \mp c \cdot \sqrt{\frac{\hat{p}_A(1-\hat{p}_A)}{n_A} + \frac{\hat{p}_B(1-\hat{p}_B)}{n_B}}\right) = (-0.152, -0.014)$$

med $c = 1.96$ från normalfördeln. tabell

5. Regressionsmodell: $Y_j = \beta_0 + \beta_1 x_j + \varepsilon_j$

a) Hypoteser: $H_0: \beta_1 = 0$
$H_1: \beta_1 < 0$

Konfidensintervall: $I_{\beta_1} = (-\infty, \hat{\beta}_1 - c \cdot s/\sqrt{S_{xx}}) = (-\infty, 0.9)$

med $c = -2.764$ ur $t(10)$-tabell

Test: $u = \frac{\hat{\beta}_1}{s/\sqrt{S_{xx}}} = \frac{-6.007}{1.844} = -3.26$

$H_0$ förkastas om $u < c$ vilket är uppfyllt.

b) $R^2 = \frac{SS_R}{SS_T} = \frac{SS_R}{SS_R + SS_E} = \frac{929.39}{875.53 + 929.39} \approx 0.5149$

c) (a) $\longleftrightarrow$ (iii)
(b) $\longleftrightarrow$ (ii)
(c) $\longleftrightarrow$ (i)

6. a) Parade mätningar.

$x_1, \ldots, x_7$ obs av $X_1, \ldots, X_7$ vikt före diet
$y_1, \ldots, y_7$ obs av $Y_1, \ldots, Y_7$ vikt efter diet

Modell: $X_i \sim N(\mu_i, \sigma_1^2)$
$Y_i \sim N(\mu_i + \Delta, \sigma_2^2)$

Hypoteser för test av viktminskning: $H_0 : \Delta = 0$
$H_1 : \Delta < 0$

$z_i = y_i - x_i$ obs av $Z_i \sim N(\Delta, \underbrace{\sigma_1^2 + \sigma_2^2}_{\sigma_Z^2})$

$\bar{z} = -0.0857$ och $s_z = 3.063$

Konfidensintervall: $I_\Delta = (-\infty, \bar{z} - c \cdot s/\sqrt{7}) = (-\infty, 2.16)$
med $c = -1.943$ från $t(6)$-tabell

Test: $u = \dfrac{\bar{z}}{s/\sqrt{7}} = -0.074 \Rightarrow u \not< c$

$H_0$ kan inte förkastas. Vi har inte kunnat visa att dieten är effektiv för viktnedgång.

b) I kursen har vi använt F-test för att jämföra varianser $H_0: \sigma_1^2 = \sigma_2^2$ mot $H_1: \sigma_1^2 \neq \sigma_2^2$
samt i regressionsanalys för att $\sigma_1^2 > \sigma_2^2$
pröva $H_0$: förklaringsvariabler meningslösa $\sigma_1^2 < \sigma_2^2$
$H_1$: minst en förklaringsvariabel gör nytta.

F-fördelningen uppkommer eftersom vi har två stokastiska variabler med $\chi^2$-fördeln. som vi tittar på.

Ex: $\dfrac{(n-1)S_1^2}{\sigma_1^2} \sim \chi^2(n-1)$ och $\dfrac{(m-1)S_2^2}{\sigma_2^2} \sim \chi^2(m-1)$

$\Rightarrow V = \dfrac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F(n-1, m-1)$