LÖSNINGSFÖRSLAG OMTENTAMEN 2012-04-16

Uppgift 1

a) $\hat{y} = b_0 + b_1 x$

$r_{xy} = -0.50$

Vid enkel linjär regression är förklaringsgraden R^2:

$$R^2 = r_{xy}^2 = (-0.50)^2 = 0.25$$

$\Longrightarrow$ C.

b) Antal sätt att dra 4 kort från 52 (totala utfallsrummet):

$$\binom{52}{4} = \frac{52!}{48!\,4!} = \frac{52 \cdot 51 \cdot 50 \cdot 49}{4 \cdot 3 \cdot 2 \cdot 1} = 270725$$

Antal sätt att dra 2 ess från 4: $\binom{4}{2} = \frac{4!}{2!\,2!} = 6$

Antal sätt att dra 2 icke-ess från 48:

$$\binom{48}{2} = \frac{48!}{46!\,2!} = \frac{48 \cdot 47}{2 \cdot 1} = 1128$$

Sannolikheten att 2 av 4 dragna är ett ess:

$$\frac{6 \cdot 1128}{270725} = 0.025 \Longrightarrow \text{A.}$$

c) Låt P = andel felbokförda fakturor

$n = 100 \qquad \hat{p} = \frac{7}{100} = 0.07$

95% konfidensintervall: $\hat{p} \pm \underbrace{z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}}_{\text{felmarginal}}$

Felmarginal: $1.96\sqrt{\frac{0.07(0.93)}{100}} = 0.05 \Longrightarrow$ C.

Notera att svaret hade blivit detsamma med ändlighetskorrektion då $\left(1 - \frac{n}{N}\right) \approx 1$.

Uppgift 2.

Låt R = ser reklam $\qquad K$ = köper produkt

$\bar{R}$ = ser ej reklam $\qquad \bar{K}$ = köper ej produkt

$P(R) = 0.40$

$P(\bar{R}) = 0.60$

$P(K|R) = 0.20$

$P(K|\bar{R}) = 0.10$

a) Vi söker $P(K)$.

Satsen om total sannolikhet ger att:

$$P(K) = P(K \cap R) + P(K \cap \bar{R})$$
$$= P(K|R)P(R) + P(K|\bar{R})P(\bar{R})$$
$$= 0.20 \cdot 0.40 + 0.10 \cdot 0.60 = 0.08 + 0.06 = 0.14$$

$\Longrightarrow$ E.

b) Vi söker $P(\bar{R} \cap \bar{K})$.

Ställ upp korstabell:

	R	$\bar{R}$	
K	0.08	0.06	0.14
$\bar{K}$	0.32	0.54	0.86
	0.40	0.60	1.00

Från korstabell ser vi att $P(\bar{R} \cap \bar{K}) = 0.54 \Longrightarrow$ B.

Uppgift 3.

$X = \#$ rätt för Max $\quad X \sim Bin(n=5, p=0.5)$

$Y = \#$ rätt för My $\quad Y \sim Bin(n=5, p=0.8)$

a) $E(X) = n \cdot p = 5 \cdot 0.5 = 2.5$

$E(Y) = n \cdot p = 5 \cdot 0.8 = 4$

$\Rightarrow$ B.

b) $D = Y - X$ $\quad$ Y & X är oberoende så $Cov(X,Y) = 0$

$E(D) = E(Y - X) = E(Y) - E(X) = 4 - 2.5 = 1.5$

$Var(D) = Var(Y - X) = Var(Y) + Var(X)$

$Var(X) = n \cdot p(1-p) = 5 \cdot 0.5 \cdot 0.5 = 1.25$

$Var(Y) = n \cdot p(1-p) = 5 \cdot 0.8 \cdot 0.2 = 0.8$

$Var(D) = 1.25 + 0.8 = 2.05$

$\Rightarrow$ B.

c) $P(X=3) = \binom{n}{x} p^x (1-p)^{n-x} = \binom{5}{3} 0.5^3 0.5^2 = \frac{5!}{3!\,2!} 0.5^5 = 10 \cdot 0.5^5$

$= 0.3125$

Alternativt slå upp i Tabell 2.

$\Rightarrow$ E.

d) $P(Y \geq 3) = 1 - P(Y \leq 2) = 1 - [P(Y=0) + P(Y=1) + P(Y=2)]$

$P(Y=0) = \binom{5}{0} 0.8^0 (0.2)^5 = 0.2^5 = 0.00032$

$P(Y=1) = \binom{5}{1} 0.8^1 (0.2)^4 = 5 \cdot 0.8 \cdot 0.2^4 = 0.00640$

$P(Y=2) = \binom{5}{2} 0.8^2 (0.2)^3 = \frac{5!}{2!\,3!} 0.8^2 (0.2)^3 = 10 \cdot 0.8^2 0.2^3 = 0.05120$

$P(Y \geq 3) = 1 - [0.00032 + 0.0064 + 0.0512] \approx 0.9421 \Rightarrow$ E.

Alternativt låt $Y = \#$ fel som My får, $Y \sim Bin(n=5, p=0.2)$ och slå upp $P(Y \leq 2)$ i Tabell 3.

Uppgift 4

a) Testvariabel: $\chi^2 = \sum_i \sum_j \frac{(O_{ij}-E_{ij})^2}{E_{ij}}$ där $E_{ij} = \frac{R_i C_j}{n}$

	Skift 1	Skift 2	R_i
Maskin 1	52 (49.5)	47 (49.5)	99
Maskin 2	36 (32)	28 (32)	64
Maskin 3	12 (18.5)	25 (18.5)	37
C_j	100	100	200 = n

Exempel på beräkning av förväntade frekvenser:
(anges inom parentes i tabellen ovan)

$E_{11} = \frac{99 \cdot 100}{200} = 49.5$ $E_{12} = \frac{99 \cdot 100}{200} = 49.5$

$E_{21} = \frac{64 \cdot 100}{200} = 32$ $E_{22} = \frac{64 \cdot 100}{200} = 32$ o.s.v.

Observerat värde på χ^2:

$$\chi^2_{obs} = \frac{(52-49.5)^2}{49.5} + \frac{(47-49.5)^2}{49.5} + \frac{(36-32)^2}{32} + \frac{(28-32)^2}{32}$$
$$+ \frac{(12-18.5)^2}{18.5} + \frac{(25-18.5)^2}{18.5} =$$

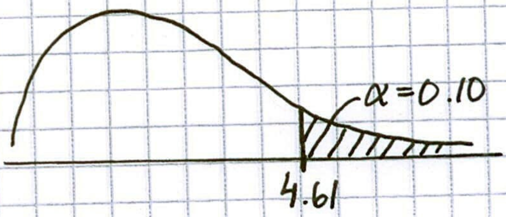
$$= 0.126 + 0.126 + 0.5 + 0.5 + 2.284 + 2.284 = 5.820 \Rightarrow A.$$

b) # frihetsgrader: $(r-1)(c-1) = (3-1)(2-1) = 2$

$\alpha = 0.10$ $\chi^2_{2,0.10} = 4.61$

Förkasta H_0 om $\chi^2_{obs} > 4.61$.

$\Rightarrow$ D.



Uppgift 5.

Låt X = betalningstiden i veckor

$X \sim N(\mu, \sigma)$ $\quad \mu = 5.8$, $\sigma = 2.4$

Urval: $n = 16$, $\bar{x} = 4.9$

a) Vi ska testa om betalningstiden har minskat dvs om $\mu < 5.8$.

$H_0: \mu = 5.8$

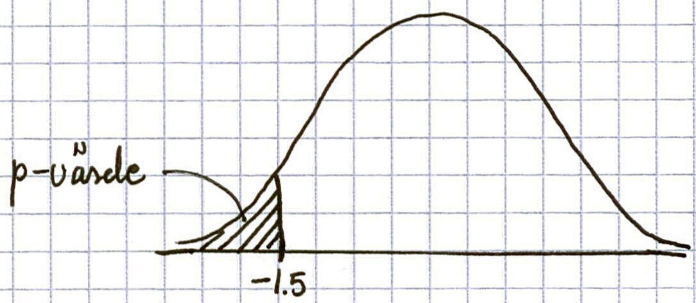
$H_1: \mu < 5.8 \implies$ A.

b) Testvariabel: $n < 30$, σ^2 känd, normalfördelad population

$$Z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

Observation: $z_{obs} = \dfrac{4.9 - 5.8}{2.4/\sqrt{16}} = -1.5 \implies$ B.

c) Enkelsidigt test.



p-värde:

$$P(Z < z_{obs}) = P(Z < -1.5)$$
$$= 1 - P(Z < 1.5)$$
$$= 1 - 0.9332 = 0.0668 \implies \text{C.}$$

Uppgift 6.

Låt P_x = andel kvinnor som vill resa utomlands
P_y = andel män —— " ——

X (kvinnor)	Y (män)
$n_x = 1400$	$n_y = 1100$
$\hat{p}_x = \frac{854}{1400} = 0.61$	$\hat{p}_y = \frac{572}{1100} = 0.52$

Vi ska testa om $P_x > P_y$:

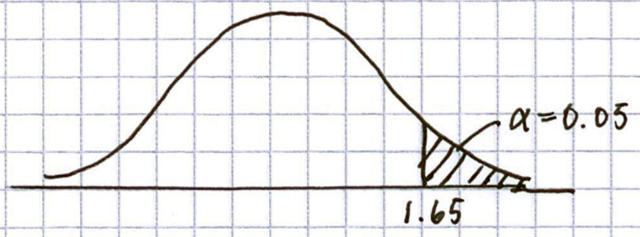
Hypoteser: $H_0: P_x - P_y = 0$ (alt. $P_x = P_y$)
$H_1: P_x - P_y > 0$ (alt. $P_x > P_y$)

Signifikansnivå: $\alpha = 0.05$

Testvariabel: $$Z = \frac{\hat{P}_x - \hat{P}_y}{\sqrt{\hat{P}_0(1-\hat{P}_0)\left(\frac{1}{n_x}+\frac{1}{n_y}\right)}} \sim N(0,1) \text{ under } H_0$$

eftersom n_x och n_y är stora stickprov som antas oberoende.
Om H_0 är sann gäller att $P_x = P_y$ och $\hat{p}_0$ är andelen i det sammanslagna stickprovet: $\hat{p}_0 = \frac{n_x \hat{P}_x + n_y \hat{P}_y}{n_x + n_y}$

Beslutsregel: Enkelsidigt test.
Förkasta H_0 om $Z_{obs} > 1.65$.



→

→ Uppgift 6 forts.

Observation:

$$\hat{p}_0 = \frac{n_x\hat{p}_x + n_y\hat{p}_y}{n_x+n_y} = \frac{1400(0.61)+1100(0.52)}{2500} = 0.5704$$

$$z_{obs} = \frac{\hat{p}_x - \hat{p}_y}{\sqrt{\hat{p}_0(1-\hat{p}_0)\left(\frac{1}{n_x}+\frac{1}{n_y}\right)}} = \frac{0.61-0.52}{\sqrt{0.5704(1-0.5704)\left(\frac{1}{1400}+\frac{1}{1100}\right)}}$$

$$= \frac{0.09}{0.01994} \approx 4.51$$

Slutsats: H_0 förkastas då $z_{obs} > 1.65$.
Vi har stöd på 5% signifikansnivå att kvinnor är mer benägna än män att vilja resa utomlands.

Notera att du även kan definiera
p_x = andel män $\quad$ p_y = andel kvinnor
Då blir hypoteserna istället:
$H_0: \ p_x - p_y = 0$
$H_1: \ p_x - p_y < 0$
och vi förkastar i vänstra svansen av fördelningen då
$z_{obs} = -4.5 < z_{krit} = -1.65$.