LÖSNINGSFÖRSLAG TENTAMEN 2012-03-16

Uppgift 1

a) $A \sim N(\mu = 50, \sigma = 3)$

$B \sim N(\mu = 50, \sigma = 7)$

$C \sim N(\mu = 50, \sigma = 5)$

Korrekt rangordning från lägsta till högsta σ: A, C, B

$\Rightarrow$ D.

b) p-värde = 0.03, α = signifikansnivå

H_0 förkastas när $\alpha >$ p-värde $\Rightarrow$ E.

c) Låt P = andel ordrar som <u>inte</u> levereras inom 3 dagar

Urval $n = 100$

i urval som levererades inom 3 dagar = 87

i urval som inte levererades inom 3 dagar = 100 − 87 = 13

$$\hat{p} = \frac{13}{100} = 0.13$$

95% konfidensintervall för P:

$$\hat{p} \pm z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$0.13 \pm 1.96 \sqrt{\frac{0.13(1-0.13)}{100}}$$

$$0.13 \pm 0.066$$

$$[0.064 ; 0.196] \Rightarrow B.$$

d) Låt C = läser City och M = läser Metro

$P(M) = 0.52$, $P(C) = 0.34$, $P(M \cap C) = 0.13$

Vi söker $P(C|M) = \frac{P(M \cap C)}{P(M)} = \frac{0.13}{0.52} = 0.25 \Rightarrow$ B.

Uppgift 2

$X = \#$ som röstar för $\quad X \sim \text{Binomial}(n=7,\ p=0.3)$

$Y = \#$ som röstar mot $\quad Y \sim \text{Binomial}(n=7,\ p=0.7)$

a) $P(X=4) = \binom{n}{x} p^x (1-p)^{n-x}$

$$= \binom{7}{4} 0.3^4 (1-0.3)^{7-4}$$

$$= \frac{7!}{(7-4)!\,4!} 0.3^4\, 0.7^3$$

$$= \frac{7 \cdot 6 \cdot 5}{3 \cdot 2 \cdot 1} 0.3^4\, 0.7^3 = 0.09724$$

$$\approx 0.097 \Rightarrow E.$$

(alternativt slå upp i Tabell 2)

b) $P(Y \geq 3) = 1 - P(Y \leq 2)$

$$= 1 - [P(Y=0) + P(Y=1) + P(Y=2)]$$

$$P(Y=0) = \binom{7}{0} 0.7^0 (1-0.7)^{7-0} = 0.3^7 = 0.0002$$

$$P(Y=1) = \binom{7}{1} 0.7^1 (1-0.7)^{7-1} = 7 \cdot 0.7 \cdot 0.3^6 = 0.0036$$

$$P(Y=2) = \binom{7}{2} 0.7^2 (1-0.7)^{7-2} = 21 \cdot 0.7^2\, 0.3^5 = 0.0250$$

$$\Rightarrow 1 - P(Y \leq 2) = 1 - [P(Y=0) + P(Y=1) + P(Y=2)]$$

$$= 1 - [0.0002 + 0.0036 + 0.025]$$

$$= 1 - 0.0288 = 0.9712$$

$$\approx 0.971 \Rightarrow A.$$

Alternativt: slå upp $P(X \leq 4)$ i Tabell 2 eller Tabell 3.

Uppgift 3.

Låt A = företag A | Marknadsandelar: A = 45%
B = företag B | B = 40%
Ö = övriga företag | Ö = 15%

Hypotesprövning: H_0: $P(A)=0.45$, $P(B)=0.40$, $P(\ddot{O})=0.15$

H_1: ej sannolikhetsfördelningen i H_0

$\Rightarrow$ χ^2-test av anpassning ("Goodness of fit")

a) Testvariabel:

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} \sim \chi^2(k-1) \Rightarrow C.$$

b) Urval $n=200$

Företag	i	O_i	P_i	$E_i = n \cdot P_i$
A	1	102	0.45	$200 \cdot 0.45 = 90$
B	2	82	0.40	$200 \cdot 0.40 = 80$
Ö	3	16	0.15	$200 \cdot 0.15 = 30$
Σ:		200	1.00	200

Observation: $$\chi^2_{obs} = \frac{(102-90)^2}{90} + \frac{(82-80)^2}{80} + \frac{(16-30)^2}{30}$$

$$= 1.60 + 0.05 + 6.53 = 8.18 \Rightarrow A.$$

c) $\alpha = 0.05$

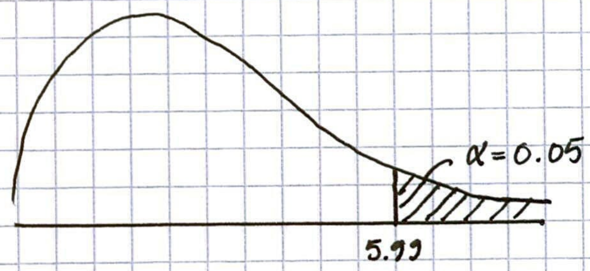
frihetsgrader $= k-1 = 3-1 = 2$

$\chi^2_{2,0.05} = 5.99$

Förkasta H_0 om $\chi^2_{obs} > 5.99$

$\chi^2_{obs} = 8.18 \Rightarrow$ förkasta H_0

$\Rightarrow B.$



Uppgift 4.

Urval: $n = 100$

Urvalsmedelvärde: $\bar{x} = 299$

Urvalsstandardavvikelse: $s = 40$

Skatteverket vill testa om företagets redovisade värde $\mu = 288$ är för lågt dvs om $\mu > 288$

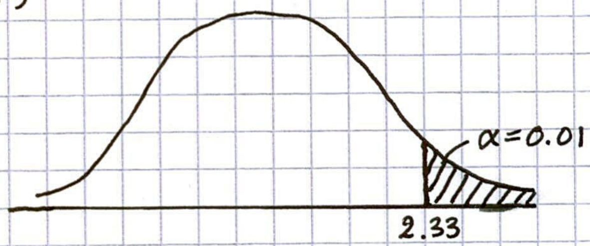
a) $H_0: \mu = 288$

$H_1: \mu > 288 \Longrightarrow$ E.

b) $\alpha = 0.01$

Testvariabel: $n > 30$, σ^2 okänd

$$\Longrightarrow Z_1 = \frac{\bar{x} - \mu}{s/\sqrt{n}} \sim N(0,1)$$



Enkelsidigt test: $z_\alpha = z_{0.01} = 2.33$

Förkasta H_0 om $z_{obs} > 2.33$

$\Longrightarrow$ C.

c) Observation:

$$z_{obs} = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{299 - 288}{40/\sqrt{100}} = 2.75 \Longrightarrow A.$$

Uppgift 5.

Simultana sannolikhetsfördelningen för X och Y:

		x: 100	x: 500	P(y)
y	100	0.40	0.20	0.60
y	500	0.30	0.10	0.40
	P(x)	0.70	0.30	1.00

P(y) ← marginalfördelningen för Y

P(x) ↑ marginalfördelningen för X

a) $Cov(X,Y) = \sum_{\text{alla } x} \sum_{\text{alla } y} x \cdot y \, P(x,y) - \mu_x \mu_y$

där $\mu_x = E(X)$ och $\mu_y = E(Y)$

Beräkna $E(X)$ och $E(Y)$:

x	P(x)	x·P(x)
100	0.70	70
500	0.30	150

$\sum: 220 = E(X)$

y	P(y)	y·P(y)
100	0.60	60
500	0.40	200

$\sum: 260 = E(Y)$

Beräkna $\sum_{\text{alla } x} \sum_{\text{alla } y} x \cdot y \, P(x,y)$:

x, y	P(x,y)	x·y·P(x,y)
100, 100	0.40	4000
100, 500	0.30	15000
500, 100	0.20	10000
500, 500	0.10	25000

$54000 = \sum \sum x \cdot y \, P(x,y)$

$Cov(X,Y) = 54000 - (220)(260) = -3200$

$\Longrightarrow$ D.

→ Uppgift 5 forts.

b) $Z = 5X + 10Y$

$E(Z) = E(5X + 10Y) = 5E(X) + 10E(Y) = 5(220) + 10(260) = \underline{\underline{3700}}$

$Var(Z) = Var(5X + 10Y)$

OBS! X och Y är ej oberoende eftersom $Cov(X,Y) > 0$

Använd räkneregel: $Var(aX + bY) = a^2 Var(X) + b^2 Var(Y) + 2ab\,Cov(X,Y)$

I vårat fall är alltså $a = 5$ och $b = 10$

Beräkna variansen för X och Y:

x	x^2	$P(x)$	$x^2 \cdot P(x)$
100	100000	0.70	7000
500	250000	0.30	75000

$\sum : E(X^2) = 82000$

$Var(X) = E(X^2) - E(X)^2 = 82000 - (220)^2 = 33600$

y	y^2	$P(y)$	$y^2 \cdot P(y)$
100	10000	0.60	6000
500	250000	0.40	100000

$\sum : E(Y^2) = 106000$

$Var(Y) = E(Y^2) - E(Y)^2 = 106000 - (260)^2 = 38400$

$$Var(Z) = Var(5X + 10Y) = 5^2 Var(X) + 10^2 Var(Y) + 2(5)(10)(-3200)$$
$$= 25(33600) + 100(38400) + 100(-3200)$$
$$= \underline{\underline{4360000}} \Rightarrow D.$$

Notera att denna uppgift går att lösa även om du ej har löst $Cov(X,Y)$ i (a). Givet korrekt beräkning av $E(Z)$, $Var(X)$ och $Var(Y)$ kan du ställa upp ekvationer för att lösa ut $Cov(X,Y)$ där $Var(Z)$ då är 2 möjliga alternativ (givet $E(Z) = 3700$)
På så sätt får du även svaret till (a).

Uppgift 6.

Parvisa Observationer, $n = 5$

Låt $T1$ = tid att skriva med Tangentbord 1

$T2$ = tid att skriva med Tangentbord 2

X_D = tidsskillnaden $= T1 - T2$

Läsare	T1	T2	X_D
1	55	56	-1
2	38	34	4
3	65	50	15
4	57	45	12
5	35	40	-5

$\sum_i : 25$

Urvalsmedelvärde: $\overline{X}_D = \dfrac{\sum_i X_D}{n} = \dfrac{25}{5} = 5$

Urvalsvarians: $S_D^2 = \dfrac{\sum_i (X_D - \overline{X}_D)^2}{n-1}$

$$= \frac{(-1-5)^2 + (4-5)^2 + (15-5)^2 + (12-5)^2 + (-5-5)^2}{5-1}$$

$$= \frac{286}{4} = 71.5$$

Urvalsstandardavvikelse: $S = \sqrt{S^2} = \sqrt{71.5} \approx 8.456$

Om det går fortare att skriva med Tangentbord 2 så är $T1 > T2$ dvs $T1 - T2 > 0$

Hypotesprövning: $H_0: \mu_D = 0$ (dvs ingen skillnad, $T1 = T2$)

$H_1: \mu_D > 0$ (dvs positiv skillnad, $T1 > T2$)

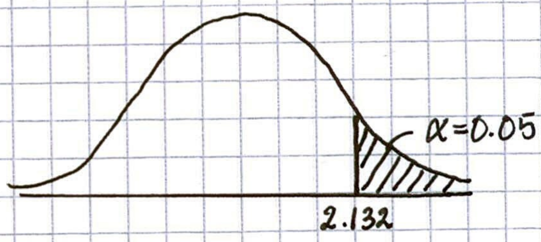
Signifikansnivå: $\alpha = 0.05$

$\longrightarrow$

→ Uppgift 6 forts.

Testvariabel: $n < 30$, σ_D^2 okänd, normalfördelad population

$$\Rightarrow t = \frac{\bar{x}_D - \mu_D}{s_D/\sqrt{n}} \sim t(n-1)$$



Beslutsregel: Enkelsidigt test

frihetsgrader $= n-1 = 4$

$t_{4,0.05} = 2.132$

Förkasta H_0 om $t_{obs} > 2.132$

Observation:

$$t_{obs} = \frac{\bar{x}_D - \mu_D}{s_D/\sqrt{n}} = \frac{5-0}{8.456/\sqrt{5}} \approx 1.322$$

Slutsats: Förkasta ej H_0 eftersom $t_{obs} = 1.322 < 2.132$.
Vi har ej kunnat påvisa att det går att skriva fortare med Tangentbord 2 på 5% signifikansnivå.