Norges teknisk–
naturvitenskapelige universitet
Institutt for matematiske fag

LØSNINGSFORSLAG
EKSAMEN I ST1201/ST6201 STATISTISKE METODER
Fredag 14. desember 2012
Tid: 09:00–13:00

Oppgave 1

Let $X_1, ..., X_{100}$ be a random sample from a normal distribution with unknown expectation μ and variance $\sigma^2 = 25$. The hypothesis $H_0 : \mu = 0$ is tested against $H_1 : \mu > 0$ (H_0 is rejected for large values of $\bar{X}$). For $\mu = 1$ the power of the test is $1 - \beta(1) = 0.5$.

a) What does the significance level α equal?

Solution. The power function (probability to reject H_0) is

$$1 - \beta(\mu) = \Phi\left(\frac{\sqrt{n}}{\sigma}\mu - z_\alpha\right).$$

In our case

$$1 - \beta(1) = \Phi(2 - z_\alpha) = 0.5$$

i.e. $z_\alpha = 2$, $\alpha = 0.0228$.

b) Find the power $1 - \beta(2)$ for $\mu = 2$.

Solution.

$$1 - \beta(2) = \Phi(4 - z_\alpha) = \Phi(2) = 0.9772.$$

Oppgave 2

Two independent samples of sizes $n = 200$ and $m = 240$ are taken from normal distributions with unknown expectations μ_X, μ_Y and known variances $\sigma_X^2 = 1$ and $\sigma_Y^2 = 1.2$, respectively. $H_0 : \mu_X = \mu_Y$ is being tested against $H_1 : \mu_X \neq \mu_Y$.

a) Find the $P-$value if observed sample means are $\bar{x} = 2.1$ and $\bar{y} = 2.0$.

Solution. Under H_0 the test statistic of the test

$$\frac{\bar{X} - \bar{Y}}{\sqrt{\sigma_X^2/n + \sigma_Y^2/m}}$$

has the standard normal distribution, therefore the $P-$value is

$$p(x, y) = P_{\mu_X = \mu_Y}\left(\left|\frac{\bar{X} - \bar{Y}}{\sqrt{\sigma_X^2/n + \sigma_Y^2/m}}\right| \geq \frac{\bar{x} - \bar{y}}{\sqrt{\sigma_X^2/n + \sigma_Y^2/m}}\right) =$$

$$= 2\Phi(-1) \approx 0.32.$$

Oppgave 3

The following result is well-known.

A. If the random vector (X, Y) has a bivariate normal distribution, and X, Y are uncorrelated (the correlation coefficient $\rho(X, Y) = 0$), then X and Y are independent.

Consider the following example. Let X and T be independent random variables, X has the standard normal distribution, T takes on two values -1 and 1, each with probability 1/2. Let $Y = TX$.

a) Show that Y has a normal distribution and therefore both components X and Y of the bivariate random vector (X, Y) are normal.

Solution. Using the total probability formula, find the cumulative distribution function $F_Y(y)$ of Y:

$$F_Y(y) = P(Y \leq y) = P(Y \leq y|T = -1)P(T = -1) + P(Y \leq y|T = 1)P(T = 1) =$$

$$= P(TX \leq y|T = -1)P(T = -1) + P(TX \leq y|T = 1)P(T = 1) =$$

$$= P(-X \leq y|T = -1)P(T = -1) + P(X \leq y|T = 1)P(T = 1) =$$

$$= P(-X \leq y)P(T = -1) + P(X \leq y)P(T = 1) =$$

$$= \frac{1}{2}[P(X \geq -y) + P(X \leq y)] = P(X \leq y) = F_X(y)$$

i.e. Y has the same distribution as X, the standard normal distribution.

b) Show that $\rho(X,Y) = 0$ but X and Y are dependent.

Solution. We have

$$P(|X| > 1) = 2\frac{1}{\sqrt{2\pi}}\int_1^\infty e^{-u^2/2}du > 0,$$
$$P(|Y| < 1) = \frac{1}{\sqrt{2\pi}}\int_{-1}^1 e^{-u^2/2}du > 0.$$

But, since $|X| = |Y|$,

$$P(|X| > 1, |Y| < 1) = 0$$

therefore

$$P(|X| > 1, |Y| < 1) \neq P(|X| > 1)P(|Y| < 1)$$

i.e. X and Y are dependent. On the other hand,

$$Cov(X,Y) = E(XY) = E(TX^2) = ET \cdot EX^2 = 0$$

(since $ET = 0$).

c) Explain, why the example of this problem is not in contradiction with proposition A.

Solution. Normality of both components of a bivariate random vector does not imply (generally speaking) that the vector has a bivariate normal distribution.

Oppgave 4

A researcher would like to find out (using ANOVA technique) whether a woman's name affects hers weight. The data (weights of 12 women) are given in the table.

Anna	Elsa	Julia
67	53	63
48	61	69
50	72	51
52	75	54

a) Show the ANOVA table (without "$P-$value"-column).

Solution.

Source	df	SS	MS	F
Treatment	2	243	121.5	1.48
Error	9	738	82	
Total	11	981		

b) Test whether the differences among the average weights are statistically significant. The significance level $\alpha = 0.05$.

Solution. Since

$$\text{the observed } F = 1.48 < 4.26 = F_{1-\alpha,2,9},$$

H_0 (all three expectations are equal) is not rejected.

Oppgave 5

0.57	0.84	0.61	0.39	0.42	0.71	0.28	0.32
0.63	0.51	0.48	0.82	0.69	0.77	0.53	0.56

The data, presented in the Table, are 16 independent observations from a continuous, symmetric (about the unknown expectation μ) distribution. Test the hyposesis $H_0 : \mu = 0.5$ versus $H_1 : \mu > 0.5$ (significance level $\alpha = 0.05$),

a) using the large-sample sign test;

Solution. Let X be the number of observations greater than 0.5. The large-sample sign test: H_0 is rejected if

$$Z = \frac{X - n/2}{\sqrt{n/4}} \geq z_\alpha.$$

In our case ($X = 11$, $n = 16$)

$$Z = 1.5 < 1.645 = z_\alpha$$

therefore H_0 is not rejected.

b) using the large-sample Wilcoxon signed rank test.

Solution. The large-sample Wilcoxon signed rank test: H_0 is rejected if

$$z = \frac{w - n(n+1)/4}{\sqrt{n(n+1)(2n+1)/24}} \geq z_\alpha,$$

where $w = \sum_{i=1}^{n} r_i z_i$, r_i – rank of $|y_i - 0.5|$, $z_i = 1$ if $y_i > 0.5$ and 0 otherwise.

In our case

i	1	2	3	4	5	6	7	8
$\|y_i - 0.5\|$	0.07	0.34	0.11	0.11	0.08	0.21	0.22	0.18
r_i	5	16	7.5	7.5	6	12	13.5	10
z_i	1	1	1	0	0	1	0	0

i	9	10	11	12	13	14	15	16
$\|y_i - 0.5\|$	0.13	0.01	0.02	0.22	0.19	0.27	0.03	0.06
r_i	9	1	2	13.5	11	15	3	4
z_i	1	1	0	1	1	1	1	1

$$w = 97$$

and

$$z = 1.5 < 1.645 = z_\alpha.$$

Thus H_0 is not rejected.