Math 472 **Statistics** **Spring 2008**

Homework Set 4 , given Fri 2/22, due Friday 2/29

Problem 4.1. Suppose one observation $X \sim N(\mu, 1)$ with unknown μ and consider hypotheses $H_0 : \mu = \mu_0$, $H_a : \mu \neq \mu_0$. It was shown that the two-sided Z-test

$$\phi(X) = \begin{cases} 1 \text{ if } T(X) > z^*_{\alpha/2} \\ 0 \text{ otherwise.} \end{cases}$$

with a test statistic $T(X) = |X - \mu_0|$, has power function

$$\begin{aligned} \gamma(\mu) &= \Pr_\mu (\phi(X) = 1) \\ &= \Phi\left(-z^*_{\alpha/2} - |\mu_0 - \mu|\right) + \Phi\left(-z^*_{\alpha/2} + |\mu_0 - \mu|\right) \\ &= \Phi\left(-z^*_{\alpha/2} - (\mu_0 - \mu)\right) + \Phi\left(-z^*_{\alpha/2} + (\mu_0 - \mu)\right). \end{aligned}$$

Show that $\gamma(\mu)$ is strictly increasing as a function of $|\mu_0 - \mu|$, for all $\alpha > 0$.

Problem 4.2. The National Center for Health Statistics reports that the systolic blood pressure for males 35 to 44 years of age has mean 128 and standard deviation 15. The medical director of a large company looks at the medical records of 72 executives in this age group and finds that the mean systolic blood pressure in this sample is $\bar{X} = 126.07$. Is this evidence that the company's executives have a different mean blood pressure from the general population?

Suppose we know that executives' blood pressures follow a Normal distribution with standard deviation $\sigma = 15$. Set up appropriate hypotheses, carry out a test and report the P-value. Evaluate the evidence at a 5% significance level.

Problem 4.3. **(MATLAB exploration).** *It is desirable here that you have access to the software and carry out the described procedures, in order to familiarize yourself with its statistics package. However for answering the question below it suffices to read the text.*

Enter the command
>> x = normrnd(0,1,1,100)
This generates 100 simulated realizations of Z, the standard normal $N(0, 1)$ random variable. They are organized as a row vector of values $x = (x_1, \ldots, x_{100})$. "In normrnd(0,1,1,100)" the first two numbers (0,1) are the parameters of the normal. In MATLAB the output is always a matrix, so the last two numbers (1,100) are the dimensions of the output matrix: it is a 1×100 matrix, i.e. a row vector.
Enter the command
>>x
This allows you to see the data, i.e. all components of the data vector x are shown.
>>m = mean(x)
This returns the sample mean $\bar{x}$ of the vector x. Enter the command
>>[h,p]=ztest(x,0.25,1)
This carries out a two a two-sided Z-test of the null hypothesis $\mu = 0.25$ (second entry in "ztest(·,·,·)" assuming that x is an i.i.d. sample from $N(\mu, \sigma^2)$ where σ is the third entry in "ztest(·,·,·)", i.e. $\sigma = 1$ our case. (Note that the standard deviation σ is the input, not the

variance σ^2). The output symbol [h,p] means that h is the the decision, 0 or 1, made with default significance level 5%. The "p" means that the P-value is also returned; an example for an output is "h =1, p = 0.0433". This means that the two-sided P-value is 0.0433, smaller than $\alpha = 0.05$ and thus decision $h = 1$ is reached (rejection of the null hypothesis).

Now let us add some more features. The command
>>[h,p,ci,zval]=ztest(x,0.25,1)
gives as an output "ci" a confidence interval at the standard 5% level, e.g. [-0.1481 0.2439]. This agrees with the decision "h=1" since 0.25 is outside the interval. Furthermore it gives the the value of the Z-statistic i.e. $Z = \sqrt{n}\left(\bar{X} - \mu_0\right)/\sigma$ which is standard normal under the null hypothesis. (Check that the software does not cheat by finding the P-value using zval and the normal table.) We can also vary the significance level α for the decision output h and ci, by e.g. setting $\alpha = 0.01$ by the command
>>[h,p,ci]=ztest(x,0.25,1,0.01).
Now our decision h is based on $\alpha = 0.01$, and so is ci. Check that ci is wider than the previous (default) one based on $\alpha = 0.05$. Check also that h always agrees with ci: if ci does not cover $\mu_0 = 0.25$ then h=1.

Question. *Suppose we wish to carry out a one sided Z-test, for alternative* $\mu < \mu_0 = 0.25$*. Matlab has a command for that, but in our version it does not work for some reason, and we don't have a normal table. How can we get the decision (call it* $h1$*) of the one sided test for the hypothesis* $\mu < \mu_0 = 0.25$*, at level* $\alpha = 0.01$*, using only one or several of the above commands ? There are several possibilities; give at least two of them: one which includes use of the output p and one which includes use of the output ci. (Don't program h1, describe it).*