Sec. 4.6

Day 10 (2/14)

1 Comments about statistical tests

1.1 Two-sided tests

Let $X_1, X_2, \ldots, X_n$ be a random sample from a normal distribution with mean μ and standard deviation $\sigma = 5000$. Suppose that we want to test

$$H_0 : \mu = 30,000 \text{ versus } H_1 : \mu \neq 30,000$$

It makes sense that we should reject H_0 if the sample mean is too small or too large, i.e.

$$\text{Reject } H_0 \text{ in favor of } H_1 \text{ if } \overline{X} \leq h \text{ or } \overline{X} \geq k$$

The benchmarks h and k are chosen so that the test achieves the desired size

$$\alpha = P_{H_0}\left(\overline{X} \leq h \text{ or } \overline{X} \geq k\right) = P_{H_0}\left(\overline{X} \leq h\right) + P_{H_0}\left(\overline{X} \geq k\right)$$

Suppose we choose to allocate α equally, then

$$P_{H_0}\left(\overline{X} \leq h\right) = \alpha/2 \text{ and } P_{H_0}\left(\overline{X} \geq k\right) = \alpha/2$$

Under H_0, the statistic $(\overline{X} - 30000)/(5000/\sqrt{n})$ is $N(0,1)$. Since

$$\alpha/2 = P\left(\frac{\overline{X} - 30000}{5000/\sqrt{n}} \leq -z_{\alpha/2}\right) = P\left(\overline{X} \leq 30000 - z_{\alpha/2}5000/\sqrt{n}\right) \tag{1}$$

then $h = 30000 - z_{\alpha/2}5000/\sqrt{n}$. Similarly,

$$\alpha/2 = P\left(\frac{\overline{X} - 30000}{5000/\sqrt{n}} \geq z_{\alpha/2}\right) = P\left(\overline{X} \geq 30000 + z_{\alpha/2}5000/\sqrt{n}\right) \tag{2}$$

so that $k = 30000 + z_{\alpha/2}5000/\sqrt{n}$. For illustration, suppose $\alpha = .05$ and $n = 25$, then $h = 30000 - 1.965000/\sqrt{25} = 28040$ and $k = 30000 + 1.965000/\sqrt{25} = 31960$. We can write equations (1) and (2) together as

$$\alpha = P\left(\left|\frac{\overline{X} - 30000}{5000/\sqrt{n}}\right| \geq z_{\alpha/2}\right)$$

which gives a more compact expression for the decision rule

$$\text{Reject } H_0 \text{ in favor of } H_1 \text{ if } |\overline{X} - 30000| \geq z_{\alpha/2} 5000/\sqrt{n}$$

More generally, let $X_1, X_2, \ldots, X_n$ be a random sample from a distribution f with mean μ and standard deviation σ. A large-sample size-α test for

$$H_0 : \mu = \mu_0 \text{ versus } H_1 : \mu \neq \mu_0$$

is given by the approximate decision rule

$$\text{Reject } H_0 \text{ in favor of } H_1 \text{ if } \left|\frac{\overline{X} - \mu_0}{S/\sqrt{n}}\right| \geq z_{\alpha/2}$$

This is equation (4.6.4) in your textbook.

1.2 Power function of the two-sided test

Now, suppose that the true value of μ is in H_1. Recall that the power of the test is the probability that the test rejects at that value of μ. From the critical regions in equations (1) and (2), the power function is

$$\begin{aligned}
\gamma(\mu) &= P_\mu\left(\overline{X} \leq \mu_0 - z_{\alpha/2}\sigma/\sqrt{n}\right) + P_\mu\left(\overline{X} \geq \mu_0 + z_{\alpha/2}\sigma/\sqrt{n}\right) \\
&= P_\mu\left(\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \leq \frac{\mu_0 - \mu - z_{\alpha/2}\sigma/\sqrt{n}}{\sigma/\sqrt{n}}\right) + P_\mu\left(\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \geq \frac{\mu_0 - \mu + z_{\alpha/2}\sigma/\sqrt{n}}{\sigma/\sqrt{n}}\right) \\
&= \Phi\left(\frac{\sqrt{n}(\mu_0 - \mu)}{\sigma} - z_{\alpha/2}\right) + 1 - \Phi\left(\frac{\sqrt{n}(\mu_0 - \mu)}{\sigma} + z_{\alpha/2}\right)
\end{aligned}$$

where $\Phi(\cdot)$ is the standard normal cdf.

1.3 Relationship between two-sided test and confidence interval

Example:

$$\begin{aligned}
\text{Reject } H_0 : \mu = 30000 \text{ if and only if } & \left|\frac{\overline{X} - 30000}{\sigma/\sqrt{n}}\right| > z_{\alpha/2} \\
\text{if and only if } & \left|\overline{X} - 30000\right| > z_{\alpha/2}\sigma/\sqrt{n} \\
\text{if and only if } & \overline{X} \text{ is farther than } z_{\alpha/2}\sigma/\sqrt{n} \text{ from } 30000 \\
\text{if and only if } & \overline{X} \pm z_{\alpha/2}\sigma/\sqrt{n} \text{ does not contain } 30000
\end{aligned}$$

In general, suppose we test

$$H_0 : \theta = \theta_0 \text{ versus } H_1 : \theta \neq \theta_0$$

then using the size-α test

$$\text{Reject } H_0 \text{ if } \left|\frac{\hat{\theta} - \theta_0}{\text{SE}}\right| > c_\alpha$$

is equivalent to

$$\text{Reject } H_0 \text{ if the (1-}\alpha\text{)100\% confidence interval } \hat{\theta} \pm c_\alpha \cdot \text{SE does not contain 0.}$$

Table 1: Summary of tests

Parameter	Estimate	SE	c_α	Comment
μ	$\overline{X}$	$s/\sqrt{n}$	$z_{\alpha/2}$	Large sample
	$\overline{X}$	$s/\sqrt{n}$	$t_{\alpha/2,n-1}$	Exact but needs normal population
p	$\hat{p}$	$\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$	$z_{\alpha/2}$	Large sample
$\mu_1-\mu_2$	$\overline{X}_1-\overline{X}_2$	$\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}$	$z_{\alpha/2}$	Large sample
		$\sqrt{S_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$	$t_{\alpha/2,n_1+n_2-2}$	Needs normal, $\sigma_1^2=\sigma_2^2$
		$\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}$	$t_{\alpha/2,n^*}$	Needs normal
p_1-p_2	$\hat{p}_1-\hat{p}_2$	$\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$	$z_{\alpha/2}$	Large sample
		$\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$		For $H_0: p_1=p_2$ where $\hat{p}=\frac{X_1+X_2}{n_1+n_2}$

R simulation of Welch versus pooled-t

```
n1<-35                        # Sample size for x
n2<-25                        # Sample size for y
sigma1<-50                    # Population sd for x
sigma2<-50                    # Population sd for y
mu1<-40                       # Population mean for x
mu2<-0                        # Population mean for y
nsim<-10000                   # Number of trials
pval1<-numeric(nsim)          # Storage for p-value of pooled-t
pval2<-numeric(nsim)          # Storage for p-value of Welch-t
for(i in 1:nsim){
xsim<-rnorm(n1,mu1,sigma1)  # Generate x-data
ysim<-rnorm(n2,mu2,sigma2)  # Generate y-data
pval1[i]<-t.test(xsim,ysim,alternative="two.sided",var.equal=TRUE)$p.value
pval2[i]<-t.test(xsim,ysim,alternative="two.sided",var.equal=FALSE)$p.value
}
cbind(mean(pval1<.05),mean(pval2<.05))
```

```
# Write a function
welch_sim<-function(nsim=10000, n1=30, n2=30, mu1=0, mu2=0, sigma1=1,sigma2=1){
pval1<-numeric(nsim)        # Storage for p-value of pooled-t
pval2<-numeric(nsim)        # Storage for p-value of Welch-t
for(i in 1:nsim){
xsim<-rnorm(n1,mu1,sigma1)  # Generate x-data
ysim<-rnorm(n2,mu2,sigma2)  # Generate y-data
pval1[i]<-t.test(xsim,ysim,alternative="two.sided",var.equal=TRUE)$p.value
pval2[i]<-t.test(xsim,ysim,alternative="two.sided",var.equal=FALSE)$p.value
}                           # End of for() loop
return(c(mean(pval1<.05),mean(pval2<.05)))
}                           # End of function()

# Call the function
welch_sim(10000,35,25,mu1=0,0,50,50)
welch_sim(10000,35,25,mu1=10,0,50,50)
welch_sim(10000,35,25,mu1=20,0,50,50)
welch_sim(10000,35,25,mu1=30,0,50,50)
welch_sim(10000,35,25,mu1=40,0,50,50)
# Use unequal variance
welch_sim(10000,35,25,mu1=0,0,20,50)
welch_sim(10000,35,25,mu1=0,0,50,20)

# Use for() loop to automate
muvec<-c(0,10,20,30,40)
outmat<-matrix(rep(0,2*length(muvec)), ncol=2)
for(j in 1:length(muvec)){
outmat[j,]<-welch_sim(10000,35,25,mu1=muvec[j],0,20,50)
}
outmat
```