Assignment 2 - version 1

Georges Monette

Janurary 16, 2020

We consider testing $H_0 : \mu = 0$ vs $H_1 : \mu \neq 0$, where $X_1, X_2, ..., X_n$ iid $N(\mu, 1)$.

The usual 5% test rejects H_0 if $\left|\frac{\bar{X}-0}{1/\sqrt{n}}\right| > 1.96$.

Under H_0, $\frac{\bar{X}-0}{1/\sqrt{n}} = \sqrt{n}\bar{X} \sim N(0, 1)$ and, in general, $\sqrt{n}\bar{X} \sim N(\sqrt{n}\mu, 1)$

We are using:

```r
n <- 10000
mu_1 <- 5
```

1. If a test statistic has a continuous distribution under a simple H_0 ('simple' meaning that there is only distribution in H_0) then $Pr(p \leq \alpha) = \alpha$ for $\alpha \in (0, 1)$. So $Pr(p \leq 0.05) = 0.05$.

2. To find the probability that $p \leq 0.05$ if $\mu = \mu_1 = 5$,

$$\begin{aligned}
Pr(p \leq 0.05) &= Pr(\left|\frac{\bar{X}}{1/\sqrt{n}}\right| \geq 1.96) \\
&= Pr(\bar{X} \geq 1.96/\sqrt{n}) + Pr(\bar{X} \leq -1.96/\sqrt{n}) \\
&= Pr(\bar{X} - \mu_1 \geq 1.96/\sqrt{n} - \mu_1) + Pr(\bar{X} - \mu_1 \leq -1.96/\sqrt{n} - \mu_1) \\
&= Pr(\sqrt{n}(\bar{X} - \mu_1) \geq 1.96 - \sqrt{n}\mu_1) + Pr(\sqrt{n}(\bar{X} - \mu_1) \leq -1.96 - \sqrt{n}\mu_1) \\
&= Pr(Z \geq 1.96 - \sqrt{n}\mu_1) + Pr(Z \leq -1.96 - \sqrt{n}\mu_1)
\end{aligned}$$

We can write a function to compute the required probability:

```r
preject <- function(alpha, n, mu1) {
  crit_val <- qnorm(1 - alpha/2)
  pnorm(crit_val - sqrt(n)*mu1, lower.tail = FALSE) +
    pnorm(-crit_val - sqrt(n)*mu1)
}
preject(.05, 10000, 5)
```

```
[1] 1
```

So the probability of rejection is within 'machine epsilon' of 1.

3. The power of the test if $\mu_1 = 5$ is the answer to the previous question, i.e. 1.
4. I can't say anything about the probability that H_0 is true from this information alone.
5. Giving H_0 and $H_1 : \mu = \mu_1 = 5$ equal *a priori* probability,

   a.

$$
\begin{aligned}
Pr(H_0|p \leq 0.05) &= \frac{Pr(p \leq 0.05|H_0)Pr(H_0)}{Pr(p \leq 0.05|\mu = \mu_1)Pr(H_0) + Pr(p \leq 0.05|\mu = \mu_1)Pr(\mu = \mu_1)} \\
&= \frac{0.05 \times 0.5}{0.05 \times 0.5 + 1 \times 0.5} \\
&= \frac{0.05 \times 0.5}{0.05 \times 0.5 + 1 \times 0.5} \\
&= 0.047619
\end{aligned}
$$

   b. Letting $Z = \sqrt{n}\bar{X}$. Under H_0 $Z \sim N(0, 1)$ and, if $\mu = \mu_1$, $Z \sim N(\sqrt{n}\mu_1, 1)$. Letting $\phi(x)$ represent the standard normal density, the density for x if $\mu = \mu_1$ is $\phi(x - \sqrt{n}\mu_1)$. Now, $p = 0.049$ corresponds to $Z = \pm 1.969$ so we consider:

$$
\begin{aligned}
&Pr(H_0|Z = \pm 1.969) \\
&= \frac{(\phi(1.969) + \phi(-1.969))Pr(H_0)}{(\phi(1.969) + \phi(-1.969))Pr(H_0) + (\phi(1.969 - \sqrt{n}\mu_1) + \phi(-1.969 - \sqrt{n}\mu_1))P} \\
&= \frac{(\phi(1.969) + \phi(-1.969))Pr(H_0)}{(\phi(1.969) + \phi(-1.969))Pr(H_0) + (\phi(1.969 - 500) + \phi(-1.969 - 500))Pr(H} \\
&= \frac{(0.0574168 + 0.0574168)0.5}{(0.0574168 + 0.0574168)0.5 + (0 + 0)0.5} \\
&= 1
\end{aligned}
$$

6. Numerical solution: We need to find how small the p-value would need to be if we have evidence that would flip a prior probability for H_0 of 0.95 to a posterior probability less than 0.05. The easiest way to set up this requirement is to note the relative probability (or odds) form of Bayes rule:

$$
\frac{Pr(H_0|z)}{Pr(H_1|z)} = \frac{f(z|H_0)}{f(z|H_1)} \times \frac{Pr(H_0)}{Pr(H_1)}
$$

   To turn a prior probability for H_0 of 0.95 to a posterior probability of 0.05 corresponds to flipping prior odds of $\frac{0.95}{1-0.95} = 19$ to posterior odds of $frac0.051 - 0.05 = \frac{1}{19}$.
   Thus the likelihood ratio must be at most $1/19^2 = 1/361$.
   Since very small values are hard to visualize, we will use the log likelihood and our target will be $\ln(1/361) = -5.888878$.

   Both the likelihood ratio and the p-value are functions of the value of z, so we find a value of z to achieve the desired likelihood ratio and check what p-value it corresponds to. We can do this analytically or numerically.
   To do it numerically, we write a function to evaluate the likelihood ratio:

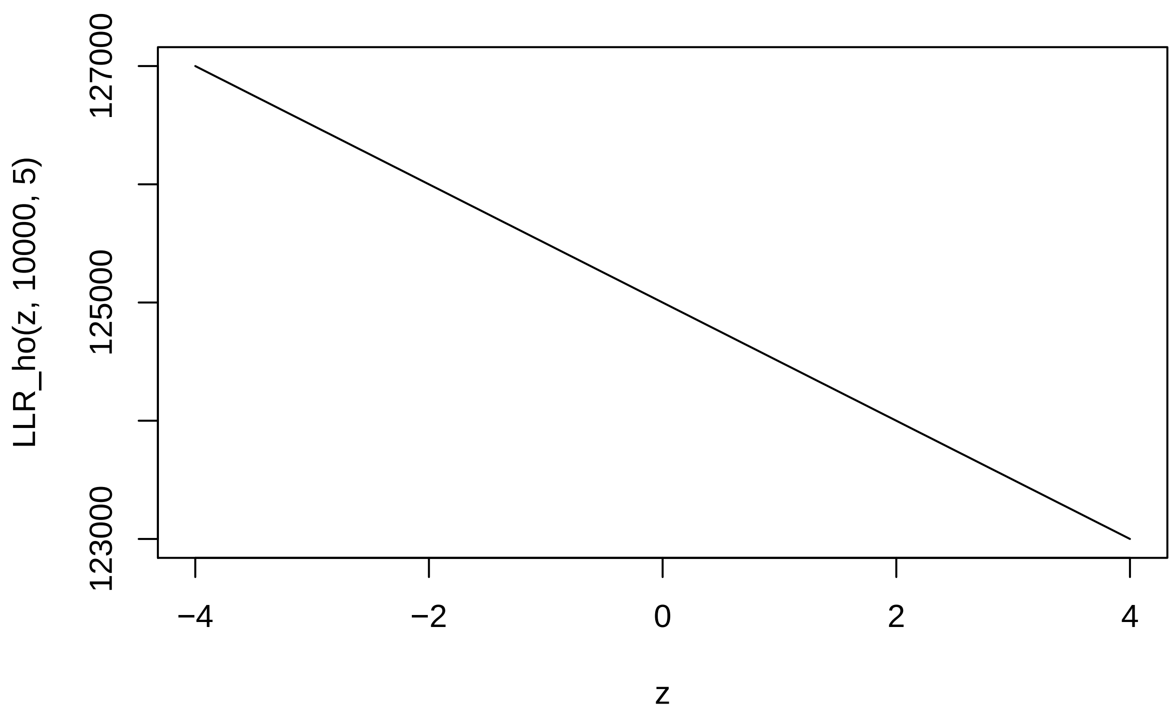
```
LLR_ho <- function(z, n, mu1) {
  dnorm(z, log = TRUE) - dnorm(z-sqrt(n)*mu1, log = TRUE)
}
```

and a function to compute the p-value:

```
pval <- function(z) {
  2 * exp(pnorm(-abs(z),log.p = TRUE))
}
```

We wish to see what value of z produces a likelihood ratio of 1/361, i.e. a log-likelihood drop of -5.888878. We can start by plotting the log-likelihood drop over a range of values of z

```
z <- seq(-4,4,.05)
plot(z, LLR_ho(z, 10000, 5), type = 'l')
```



Which doesn't take us anywhere close to -5.888878.

We can do a small regression to see what value of z would lead to rejecting H_0. If the relationship were not linear, we could use 'uniroot'.

```
fit <- lm(z ~ LLR_ho(z, 10000, 5))
(zval <- sum(coef(fit)* c(1, log(1/361))))
```

```
## [1] 250.0118
```

```r
LLR_ho(zval, 10000, 5)
```

```
## [1] -5.888878
```

The p-value to flip a prior probability of 0.95 for a null hypothesis to a posterior of 0.05, if the true value of the alternative is $\mu_1 = 5$ and $n = 10000$ is:

```r
pval(zval)
```

```
## [1] 0
```

What is the moral of this story?