T.Test

Hossein Vatani

March 26, 2016

بنام خدا

یادآوری در هر آزمون فرض ما دنبال بررسی جواب زیر هستیم

$$\begin{cases} H_0 & : \mu_0 = \mu_1 \\ H_1 & : \mu_0 \neq \mu_1 \end{cases}$$

که فرض صفر می تواند هریک از سه حالت زیر را داشته باشد

$$H_0 \begin{cases} \mu_0 = \mu_1 \\ \mu_0 \geq \mu_1 \\ \mu_0 \leq \mu_1 \end{cases}$$

آزمون یک و دو نمونه ای T-Test

تابع ()t.test در نرم افزار R امکان محاسبه هر دونوع آزمون را دارد. این تابع دارای جزءهای متفاوتی بدین شرح است

```
t.test(x, y = NULL, alternative = c("two.sided", "less", "greater"), mu = 0,
paired =
FALSE, var.equal = FALSE, conf.level = 0.95)
```

```
x,y=بردار داده های مورد نظر می باشند که برای آزمون یک جهتی نیازی به مشخص کردن
مقدار بردار دوم نیست.
alternative=جهت بررسی حالت های مختلف فرض صفر که در قسمت یادآوری آورده شده است
mu=همان عامل میانگین(میو) می باشد!
paired=جهت مشخص کردن آنکه آیا آزمون بصورت تی-زوجی است یا خیر
var.equal=آیا پراکنش(variance) دو داده(درصورت استفاده از آزمون دوجهتی) با هم ب
را برا ند؟
conf.level=سطح اطمینان مطلوب چقدر است؟
```

الف-آزمون t-test یکطرفه

مثال-بیماری خاصی از مصرف یک نوع بستنی یک کارخانه مشخص مشاهده میشد.دانشمندان میزان عامل این بیماری را در یک دسته ۹ تایی تصادفی بشرح زیر اندازه گیری کردند(واحد اندازه گیری MPN/g) 0.593 0.142 0.329 0.691 0.231 0.793 0.519 0.392 0.418 اگر میانگین آنها بیش از ۰.۳ باشد آنگاه بستنی های تولیدی بیماری زا هستند. می خواهیم بدانیم آیا بستنی های تولیدی این کارخانه بیماری زا هستند یا خیر؟ اگر μ نماد

میانگین بستنی ها باشد، جزئیات آزمون را می توان اینگونه عنوان کرد که $\begin{cases} H_0: \mu = 0.3 \\ H_1: \mu > 0.3 \end{cases}$ لذا در آزمون باید قسمت مربوط به نوع آزمون را “greater” انتخاب نماییم.

```
x = c(0.593, 0.142, 0.329, 0.691, 0.231, 0.793, 0.519, 0.392, 0.418)
t.test(x, alternative="greater", mu=0.3)
```

```
##
##  One Sample t-test
##
## data:  x
## t = 2.2051, df = 8, p-value = 0.02927
## alternative hypothesis: true mean is greater than 0.3
## 95 percent confidence interval:
##  0.3245133       Inf
## sample estimates:
## mean of x
## 0.4564444
```

همانطور که در تووضیحات نتیجه تابع در R ملاحظه می کنید، آزمون تی-استیودنت از نوع یکطرفه با سطح اطمینان ۹۵٪ می باشد. مقدار p-value به ما می گوید که باسطح اطمینان ۹۵٪ میانگین وجود عامل رمورد نظر در بستنی های تحت بررسی بیش از ۰.۳ (و برابر ۰.۴۵) می باشد.

ب-آزمون t-test دوطرفه

مثال- جهت بررسی تاثیر یک نوع دارو، دوازده نفر را به دو گروه شش نفره تقسیم کرده ایم و داروی مورد نظر را جهت مصرف به گروه اول(گروه تحت درمان) داده و به گروه دوم(گروه تحت کنترل) چیزی شبیه دارو که فاقد هرگونه اثریست و صرفا خوراکی است داده شده است. عامل مورد نظر مدت واکنش هر دوگروه بعد از مصرف برحسب میلی-ثانیه می باشد. می خواهیم بررسی کنیم که آیا بین عکس العمل دو گروه پس از مصرف اختلاف معنی داری وجود دارد یاخیر؟ در اینجا از آزمون تی-استیودنت دو طرفه استفاده می کنیم.چندین حالت مختلف را تحت بررسی قرار می دهیم. *توجه شود که در ابتدا با فرض برابری پراکنش هر دو داده باهم برابر است.* در حالت اول تنها بررسی می کنیم که آیا میانگین ها اختلاف معنی داری دارند یا خیر.

$$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{cases}$$

```
Treat = c(101, 110, 103, 93, 99, 104)
Control = c(91, 87, 99, 77, 88, 91)
t.test(Treat,Control, var.equal=TRUE)
```

```
##
##  Two Sample t-test
##
## data:  Treat and Control
## t = 3.4456, df = 10, p-value = 0.006272
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   4.534534 21.132133
## sample estimates:
## mean of x mean of y
## 101.66667  88.83333
```

مقدار p-value به ما می گوید که بین دو گروه اختلاف معنی داری وجود دارد، این بدان معنی است که استفاده از دارو یا استفاده نکردن از آن نتایج متفاوتی را بوجود می آورد.در واقع فرض صفر رد می شود. اینبار شرط آزمون فرض را بصورت زیر تغییر می دهیم. $\begin{cases} H_0 : \mu_1 \leq \mu_2 \\ H_1 : \mu_1 > \mu_2 \end{cases}$

```
t.test(Treat,Control, alternative = "greater", var.equal=TRUE)
```

```
##
##  Two Sample t-test
##
## data:  Treat and Control
## t = 3.4456, df = 10, p-value = 0.003136
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  6.082744      Inf
## sample estimates:
## mean of x mean of y
## 101.66667  88.83333
```

در این آزمون مقدار p-value می گوید که استفاده از دارو در عکس العمل بیمار موثر است یعنی فرض صفر رد می شود. و اینبار شرط آزمون فرض را بصورت زیر تغییر می دهیم. $\begin{cases} H_0 : \mu_1 \geq \mu_2 \\ H_1 : \mu_1 < \mu_2 \end{cases}$

```
t.test(Treat,Control, alternative = "less", var.equal=TRUE)
```

```
##
##  Two Sample t-test
##
## data:  Treat and Control
## t = 3.4456, df = 10, p-value = 0.9969
## alternative hypothesis: true difference in means is less than 0
## 95 percent confidence interval:
##      -Inf 19.58392
## sample estimates:
## mean of x mean of y
## 101.66667  88.83333
```

مقدار p-value به وضوح می گوید که دلیلی برای رد فرض صفر وجود ندارد یعنی نمی توان گفت که استفاده از دارو نما نسبت به استفاده از دارو موثر است!

اما اگر فرض برابری پراکنش ها را حذف کنیم چه می شود؟ $\begin{cases} H_0 : \mu_1 \leq \mu_2 \\ H_1 : \mu_1 > \mu_2 \end{cases}$

```
t.test(Treat,Control, alternative = "greater")
```

```
##
##  Welch Two Sample t-test
##
## data:  Treat and Control
## t = 3.4456, df = 9.4797, p-value = 0.003391
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  6.044949      Inf
## sample estimates:
## mean of x mean of y
## 101.66667  88.83333
```

همانطور که ملاحظه می شود بازهم دلیلی برای قبول فرض صفر وجود ندارد و با اطمینان بالایی می توان گفت که دارو در عکس العمل بیمار تاثیر قابل توجهی دارد.

ج-آزمون t-test زوجی

اگر فرد مورد نظر متعلق به هر دوگروه درمان و کنترل باشد،با وضعیتهای بیشتری جهت آزمایش مواجه خواهیم شد.برای مثال می توان نتایج تاثیر دو داروی مختلف را بررسی کرد و یا وضعیت شخص را قبل از دارو و بعد از آن بررسی کرد.در همه این موارد جهت معنی داری نتایج، امکان استفاده از آزمون تی-استیودنت دوطرفه وجود ندارد چراکه دو داده مورد بررسی مستقل از یکدیگر نیستند، لذا دراینگونه از بررسی ها از آزمون تی-استیودنت زوجی استفاده خواهیم کرد.
مثال-می خواهیم بررسی کنیم که آیا مصرف بنزین معمولی در مقابل بنزین سوپر، مزیتی در مسافت طی شده توسط خودرو دارد یا خیر؟ برای اینکار،ده ماشین را انتخاب می کنیم و بصورت اتفاقی(توسط شیریاخط) برخی را با بنزین معمولی پر و باقیمانده را با بنزین سوپر؛پس اط طی مسافت تا خالی شدن کامل مخزن

بنزین، اینبار خودروها را با نوع دیگر بنزین تکمیل کرده و مسافت طی شده بعدی را نیز ثبت می کنیم.بررسی بصورت زیر خواهد بود.

```
reg = c(16, 20, 21, 22, 23, 22, 27, 25, 27, 28)  # بنزین معمولی
prem = c(19, 22, 24, 24, 25, 25, 26, 26, 28, 32) # بنزین سوپر
t.test(prem, reg, paired=TRUE)
```

```
##
##  Paired t-test
##
## data:  prem and reg
## t = 4.4721, df = 9, p-value = 0.00155
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.9883326 3.0116674
## sample estimates:
## mean of the differences
##                       2
```

خوب !مشخص است که تفاوت معنی داری وجود دارد، اما کدامیک موثر است و تاثیر آن مورد نظر ماست؟

```
t.test(prem, reg, alternative="greater", paired=TRUE)
```

```
##
##  Paired t-test
##
## data:  prem and reg
## t = 4.4721, df = 9, p-value = 0.0007749
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  1.180207      Inf
## sample estimates:
## mean of the differences
##                       2
```