

## آزمون های آماری در متلب

### آزمون تی تک نمونه‌ای (One-sample and paired-sample t-test)

آزمون t تک نمونه‌ای زمانی مورد استفاده قرار می‌گیرد که یک نمونه از جامعه داریم و می‌خواهیم میانگین آن را با یک حالت معمول و رایج یا یک میانگین فرضی مورد انتظار مقایسه کنیم. به عبارتی دیگر در این آزمون فرض مبنی بر این است که نمونه‌ای به حجم X و میانگین m از یک جامعه داریم و می‌خواهیم بدانیم که آیا این نمونه را می‌توان با یک نمونه تصادفی از جامعه‌ای که میانگین آن u است برابر دانست؟ این آزمون شبیه آزمون Z می‌باشد. با این وجود تفاوت های ظریفی با هم دارند. آزمون Z بیانگر تفاوت میانگین نمونه با میانگین جامعه بر حسب واحد استاندارد است. از طرفی دیگر محاسبه Z مستلزم معلوم بودن انحراف استاندارد جامعه می‌باشد زیرا برای محاسبه آن انحراف استاندارد توزیع میانگین‌ها ضرورت دارد. اما در آزمون t همین عمل مقایسه میانگین نمونه با میانگین جامعه را انجام می‌دهد. اما زمانی از انحراف استاندارد جامعه اطلاعی در دست نداریم. در این مواقع در این آزمون (t) به جای انحراف استاندارد جامعه از انحراف استاندارد نمونه استفاده می‌کنیم.

مقدار این آماره از طریق فرمول زیر محاسبه می‌شود:

$$d_{t1} = \frac{d}{sd} \quad d = |m - u| \quad \text{و} \quad sd = \frac{s}{\sqrt{n}}$$

در این فرمول m میانگین نمونه یا عنصر مورد بررسی و u میانگین مورد انتظار یا فرضی می‌باشد.  
درجه آزادی:  $df = n - 1$

به مثال زیر توجه کنید. در اینجا قصد داریم میانگین ماتریس A را با میانگین فرضی 10 مقایسه کنیم. به عبارتی آیا میانگین دمای ایستگاه A با میانگین 10 درجه سانتیگراد در سطح 95 درصد اطمینان برابر است یا نه؟

$$H_0: \mu_{A7.8702} = u_{10}$$

$$H_1: \mu_{A7.8702} \neq u_{10}$$

A =	$d_{t1} = \bar{m} - u = 7.8702 - 10$
9.2888	$= -2.13$
9.5897	$sd = \frac{s}{\sqrt{n}} = \frac{2.7065}{3.16} = 0.86$
9.2505	$d_{t1(7.8702-10)} = \frac{-2.13}{0.86} = -2.48$
10.5373	
11.8020	$d_{t1} = \bar{m} - u = 7.8702 - 6 = 1.87$
7.6985	$sd = \frac{s}{\sqrt{n}} = \frac{2.7065}{3.16} = .86$
5.6571	$d_{t1(7.8702-6)} = \frac{1.87}{0.86} = 2.19$
6.9652	
3.7349	
4.1781	
$\bar{m} = 7.8702 \quad s_{\bar{x}} = 2.7065$	

با مراجعه به جداول آخر کتابهای آماری می توان مقدار خطای  $d_{t1}$  را بدست آورد که برای  $d_{t1(7.8702-10)}$  مقدار خطای  $0.345$  حاصل شده است این در حالی می باشد که برای  $d_{t1(7.8702-6)}$  چیزی حدود  $0.56$  می باشد. بنابراین برای دومی تفاوت معنی داری مشاهده نمی شود.  
برای این آزمون در متلب از دستور زیر استفاده خواهیم کرد:

>> [h, p] = ttest(x, u);

در این دستور X متغیر مورد بررسی مثلا دما، و u میانگین فرضی می باشد. با اجرای این دستور دو خروجی به نام ماتریس n و p ایجاد می گردد. در خروجی n مقدار یا برابر با یک یا صفر می باشد؛ اگر یک باشد یعنی اینکه شواهد کافی برای پذیرفتن فرض صفر وجود ندارد اگر صفر باشد یعنی شواهد کافی برای رد فرض صفر وجود ندارد. ماتریس p مقدار خطای موجود را می دهد. طبیعی است که اگر  $n=1$  و p دارای ارزشی کمتر از  $0.5$  باشد و برعکس اگر  $n=0$  و p دارای ارزشی بالاتر از  $0.5$  باشد، بدین معنی است که تفاوتی معنی داری بین میانگین وجود ندارد.

در این مثال با استفاده از این دستور قصد داریم بدانیم که آیا میانگین ماتریس A با میانگین ۶ تفاوت دارد یا نه؟ برای این منظور از آزمون تی تک نمونه ای استفاده خواهیم کرد:

A =	>> [h p]=ttest(A,6)
9.2888	h =
9.5897	0
9.2505	p =
10.5373	0.0567
11.8020	>> [h p]=ttest(A,10)
7.6985	h =
5.6571	1
6.9652	p =
3.7349	0.0345
4.1781	

اگر دقت کنید  $h=0$  حاصل شده است که بدین معنی می باشد که بین میانگین داده های دمای روزانه دهمشت با میانگین ۶ تفاوتی معنی داری در سطح ۹۵ درصد وجود ندارد. این در حالی است که خطای بالای  $0.567$  تاییدی بر گفته فوق می باشد. در مثال دوم میانگین را با ۱۰ مقایسه کردیم که نتایج نشان داد تفاوت معنی داری وجود دارد ( $h=1$  و خطای  $0.345$  تاییدی بر تفاوت میانگین ماتریس A با میانگین فرضی ۱۰ است). در صورتی که بخواهیم آماره t، درجه آزادی، و ... را مشاهده کنیم دستور تی تست تک نمونه ای را به شکل زیر بازویسی می کنیم:

>> [h, p, ci, stats] = ttest(A,6)	>> [h, p, ci, stats] = ttest(A,10)
h =	h =
0	1
p =	p =
0.0567	0.0345

<pre>ci =     5.9341     9.8063 stats =     tstat: 2.1852          df: 9          sd: 2.7065</pre>	<pre>ci =     5.9341     9.8063 stats =     tstat: -2.4885          df: 9          sd: 2.7065</pre>
--	---

در این دستور **h** همان نتایج را می‌دهد که توضیحات آن داده شده است، **ci** بیانگر حد بالا و پایین، در ماتریس **state** که خود شامل زیر ماتریس‌های **tstat** (آماره **t** را می‌دهد)، **df** (درجه آزادی) و **sd** (انحراف معیار) می‌باشد.

نکته با اجرای دستورات بالا در متلب آزمون اختلاف میانگین را به طور خودکار در سطح ۹۵ اطمینان و با مقدار خطای ۰.۰۵ / آزمون می‌کند. در صورتی که شما بخواهید اختلاف میانگین را در سطح موردنظر انتخاب کنید از دستور زیر استفاده خواهید کرد:

```
>> [h, p, ci, stats] = ttest (a, b, 'alpha', value)
```

در این دستور **a** نمونه، **b** میانگین مورد انتظار، **'alpha'** تعیین سطح مورد نظر، **value** مقدار خطای انتخابی در سطح اطمینان مختلف برای مثال برای ۹۵ درصد به جای **value** از ۰.۰۵ و ۹۹ درصد از ۰.۰۱ به شرح زیر استفاده خواهیم کرد:

```
>> [h, p, ci, stats] = ttest (a, b, 'alpha', 0.05);
```

```
>> [h, p, ci, stats] = ttest (a, b, 'alpha', 0.01);
```

با تشکر دوستکامیان