

مفاهیم آمار و طرح های آزمایشی با R

Concepts of Statistics and Experimental Designs with R

نویسنده:

حسین عسکری

۱۳۹۵

|                        |                                                   |
|------------------------|---------------------------------------------------|
| سرشناسه                | : عسکری، حسین، ۱۳۵۷ -                             |
| عنوان و نام پدیدآور    | : مفاهیم آمار و طرح های آزمایشی با R / حسین عسکری |
| مشخصات نشر             | : تهران: کتاب اطهر، ۱۳۹۵                          |
| شابک                   | : ۹۷۸-۶۰۰-۸۱۳۴-۵۵-۸                               |
| وضعیت فهرست نویسی      | : قیفا                                            |
| موضوع:                 | : آمار ریاضی - داده پردازی                        |
| موضوع                  | : طرح های آزمایشی - نرم افزار                     |
| موضوع                  | : آر (زیان برنامه نویسی کامپیوتر)                 |
| رده کنگره              | : ۱۳۹۵ ع ۴۵ / آ ۹۴ / ۴۵ / ۴۵ QA۲۷۶                |
| رده دسی                | : ۵۱۹/۵۰۲۸۵۵۱۳۳                                   |
| شماره ثبت کتابخانه ملی | : ۴۴۷۷۶۵                                          |

مفاهیم آمار و  
طرح های آزمایشی با R

### د کتاب اطهر

مؤلف: حسین عسکری  
ناظر چاپ: محمد حسین اطهر  
طراحی جلد: معصومه نشانی  
شابک: ۹۷۸-۶۰-۸۱۳۴-۵۵-۸  
نوبت چاپ: ۱۳۹۵/اول  
شمارگان: ۱۰۰۰  
قیمت: ۱۵۰۰۰ تومان



آدرس: دفتر تهران- خیابان قصرالدشت - نبش بوستان سعدی- پلاک ۵۵۵ - واحد همکف

[atharbook@elenoon.ir](mailto:atharbook@elenoon.ir)

[www.atharbook.ir](http://www.atharbook.ir)

## پیش گفتار

با حمد و سپاس خداوند متعال را به پاس تمامی الطاف و نعمت‌های بیکرانش و فراهم نمودن فرصتی برای ارائه خدمتی ناچیز به حضور دانش‌پژوهان و دانشجویان و با امید به اینکه مطالب پیش رو مورد استفاده و سودمند واقع گردد.

حسب آشنایی با نرم‌افزار R و دامنه گسترده آن در طرح‌های علمی و پژوهشی در زمینه‌های مختلف علمی و قابلیت استفاده و به‌کارگیری ساده این نرم‌افزار و همچنین ضرورت یادگیری و به‌کارگیری نرم‌افزاری مناسب و قابل اتکا در تجزیه و تحلیل داده‌های آماری، نویسنده را بر آن داشت تا تجربیات و اندوخته‌های خود را در قالب کتاب حاضر با زبانی ساده و خوانا، بر اختیار پژوهشگران قرار دهد. این کتاب چکیده‌ای از مبانی آمار و سرفصل طرح‌های آزمایشی یک به زبان R می‌باشد.

بر خود لازم میدانم از تمامی عزیزان و دوستانی که بنده را در تهیه و تدوین مطالب یاری و یا از تجربیات و دانش آن‌ها در نگارش این کتاب بهره‌مند گردیده‌ام کمال تشکر را داشته باشم. لازم به ذکر است که نظرات و راهنمایی‌های خوانندگان ارجمند در رفع نواقص و کامل‌تر کردن مطالب در چاپ‌های آتی همواره موجب سپاس و امتنان خواهد بود.

حسین عسکری

اسفند ۹۵

| صفحه | فهرست                    |
|------|--------------------------|
| ۱    | مقدمه                    |
| ۱    | نصب R                    |
| ۲    | پکیج ها                  |
| ۵    | معرفی اطلاعات به R       |
| ۶    | برداری (vector)          |
| ۷    | ادغام بردار ها           |
| ۷    | محاسبات بردار            |
| ۹    | نام گذاری اجزاء بردار    |
| ۹    | ماتریس                   |
| ۱۱   | ساختار ماتریس            |
| ۱۱   | ترانسپوز                 |
| ۱۱   | ادغام ماتریس ها          |
| ۱۲   | سایر اعمال               |
| ۱۲   | لیست                     |
| ۱۳   | خرد کردن لیست            |
| ۱۳   | نام گذاری اجزاء لیست     |
| ۱۴   | ساختار داده (Data Frame) |
| ۱۴   | ساختن Data Frame         |
| ۱۵   | آمار مقدماتی             |
| ۱۶   | داده های کیفی            |
| ۲۱   | داده های کمی             |
| ۲۱   | توزیع فراوانی            |
| ۲۳   | مقیاسهای عددی            |

|    |                              |
|----|------------------------------|
| ۲۳ | میانگین                      |
| ۲۳ | میان                         |
| ۲۳ | چارک                         |
| ۲۴ | درصد                         |
| ۲۴ | دامنه                        |
| ۲۴ | دامنه چارکی                  |
| ۲۴ | کس پلات                      |
| ۲۵ | راریتاس                      |
| ۲۵ | انحراف معیار                 |
| ۲۶ | کواریانس                     |
| ۲۶ | ضریب همبستگی                 |
| ۲۶ | گشتاور مرکزی                 |
| ۲۷ | چولگی                        |
| ۲۷ | کشیدگی                       |
| ۲۷ | پلات ها                      |
| ۳۶ | توزیع احتمالی                |
| ۳۶ | توزیع باینومیل               |
| ۳۷ | توزیع پویسون                 |
| ۳۷ | توزیع یکنواخت                |
| ۳۸ | توزیع نمایی                  |
| ۳۹ | توزیع نرمال                  |
| ۳۹ | توزیع کای اسکور توزیع        |
| ۴۰ | t استیودنت                   |
| ۴۰ | F توزیع                      |
| ۴۰ | برآورد فاصله                 |
| ۴۱ | برآورد نقطه ای میانگین جمعیت |

- ۴۱ بر آورد حدود اطمینان میانگین جمعیت با واریانس معلوم
- ۴۲ بر آورد حدود اطمینان میانگین جمعیت با واریانس نامعلوم
- ۴۳ اندازه نمونه میانگین جمعیت
- ۴۳ بر آورد نقطه ای نسبت جمعیت
- ۴۳ بر آورد فاصله نسبت جمعیت
- ۴۴ اندازه نمونه نسبت جمعیت

## ۴۴ آزمون فرض

- ۴۵ آزمون یک طرفه دامنه پایین میانگین جمعیت با واریانس معلوم
- ۴۶ آزمون یک طرفه دامنه بالای میانگین جمعیت با واریانس معلوم
- ۴۷ آزمون دو طرفه میانگین جمعیت با واریانس معلوم
- ۴۸ آزمون فرض در جوامع با واریانس نامعلوم
- ۵۱ آزمون یک طرفه دامنه پایین نسبت جمعیت
- ۵۲ آزمون یک طرفه دامنه بالای نسبت جمعیت
- ۵۳ آزمون دو طرفه نسبت جمعیت

## ۵۵ خطای نوع II

- ۵۵ آزمون یک طرفه، دامنه پایین میانگین جمعیت با واریانس معلوم
- ۵۶ آزمون یک طرفه، دامنه بالای میانگین جمعیت با واریانس معلوم
- ۵۷ آزمون دو طرفه میانگین جمعیت با واریانس معلوم
- ۵۸ آزمون یک طرفه دامنه پایین میانگین جمعیت با واریانس مشخص
- ۵۹ آزمون یک طرفه، دامنه بالای میانگین جمعیت با واریانس مشخص
- ۶۰ آزمون دو طرفه میانگین جمعیت با واریانس نامشخص

- ۶۱ مقایسات دو جامعه
- ۶۲ مقایسه میانگین در مشاهدات جفت شده
- ۶۳ مقایسه میانگین بین دو نمونه مستقل
- ۶۴ مقایسه بین نسبتهای دو جامعه

## ۶۵ شایستگی برآزش

|     |                                               |
|-----|-----------------------------------------------|
| ۶۵  | شایستگی برآزش چند جمله ای                     |
| ۶۶  | آزمون مستقل بودن                              |
| ۶۷  | رگرسیون خطی ساده                              |
| ۷۰  | برآورد معادله رگرسیون                         |
| ۷۸  | عدم برآزش                                     |
| ۷۹  | رگرسیون خطی چند گانه                          |
| ۸۱  | تجزیه واریانس                                 |
| ۸۲  | مفروضات تجزیه واریانس                         |
| ۸۵  | تجزیه واریانس یک طرفه                         |
| ۹۰  | تفسیر نتایج تجزیه واریانس یک طرفه بوسیله پلات |
| ۹۱  | تجزیه واریانس یک طرفه نامتعادل                |
| ۹۲  | مقایسات زوجی با مستقل                         |
| ۹۷  | داده های چند شاهدی                            |
| ۹۹  | طرح های آزمایشی                               |
| ۱۰۰ | طرح کاملاً تصادفی                             |
| ۱۰۰ | طرح بلوک های کامل تصادفی                      |
| ۱۰۴ | بلوک های کامل تصادفی با بیش از یک مشاهده      |
| ۱۰۸ | طرح مربع لاتین                                |
| ۱۱۲ | طرح مربع لاتین تکرار شونده                    |
| ۱۱۵ | آزمایشات فاکتوریل                             |
| ۱۲۶ | کرت های خرد شده                               |
| ۱۳۱ | تجزیه کواریانس                                |

## مقدمه:

امروزه انسان در زمینه تمام علوم زیستی، اقتصادی، اجتماعی و کشاورزی در سایه اتخاذ روشهای صحیح تحقیق توانسته است به پیشرفتهای شگرفی نائل شود. تحقیق در لغت به معنی جستجو برای کسب آگاهی است. اما از نظر علمی به مجموعه فعالیتهایی اطلاق می شود که بر اساس یک نظام فکری درست در جهت آگاهی از مجهولات، پایه ریزی شده باشد.

استفاده از نرم افزار R بنا به دلایل متعددی بر سایر نرم افزارهای آماری اولویت دارد. این نرم افزار برای تجزیه تحلیل بسیار قدرتمند است.

دلایل مطلوب بودن نرم افزار R:

- ۱- قابل دسترس و در دسترس بودن این نرم افزار برای همگان.
- ۲- بررسی مباحث کاربردی در تخصصی علم آمار و دسترسی داده های واقعی علوم دیگر.
- ۳- به روز بودن بسته های نرم افزاری (Packages) مربوط به R.
- ۴- مشترک بودن زبان برنامه نویسی آن با S-PLUS.
- ۵- تبادل اطلاعات با سایت های معتبر علمی، دانشگاه های دنیا و رفع مشکلات.
- ۶- قبول بسته های نرم افزاری از طرف سایت سازنده نرم افزار.
- ۷- امکان ساخت بسته های نرم افزاری و توابع جدید برای همگان.

## نصب R:

با دسترسی به اینترنت این امر به راحتی و سادگی امکان پذیر می باشد. ابتدا به سایت CRAN(Comprehensive R Archive Network) به آدرس <http://cran.r-project.org> رفته و با توجه به سیستم عامل دستورالعمل را جهت دانلود و نصب دنبال نمائید. باید منتظر از سایت به منظور به روز بودن نرم افزار ایده مناسب می تواند باشد و چنانچه نرم افزار شما قدیمی است بهتر است قبل از دانلود فرمت جدید، آنرا حذف نمائید.

اصطلاحات:

Rconsole: مکانی که دستورات وارد و اجرا می شوند و خروجی آنالیز ظاهر می گردد.



Editor Window: (File / New scrip) مکانی مناسب جهت نوشتن کد ها و دستورات و انتقال آنها به محیط console می باشد. مزیت اصلی آن فراهم شدن ویرایش دستورات و فرامین می باشد. Vector: مجموعه ای مشتمل بر تعدادی از اعضاء یا رشته ای از داده های پیوسته. یک مقدار ویژه به تنهایی یک وکتور با یک عضو می باشد.

data.frame: داده ها را بصورت ستون (متغیر) و با ردیف هایی که نماینده یک مشاهده (فرد) هستند، سازماندهی می کنند. به عبارتی هر ردیف نشان دهنده مقادیر متغیر ها برای هر فرد می باشد. مزیت data.frame این است که متغیر هایی از انواع مختلف را می تواند با یکدیگر ترکیب نماید. لازم به ذکر است برای اجرای برخی توابع نیاز است داده ها بصورت data frame باشند.

List: مجموعه ای متشکل از انواع متفاوت داده ها را بصورت یک مجموعه واحد تبدیل می نماید.

Matrices: داده ها را بصورت ماتریس سازمان دهی می کنند. همه اجزای ماتریس باید از یک نوع داده باشند.

Function: دستورالعمل های R که انجام دهند عمل خاص می باشند. Library: مجموعه ای از پکیج های طراحی شده برای اعمال خاص که توسط افراد مختلف در جامعه R نوشته شده اند.

### پکیج ها:

در نرم افزار R به طور پیش فرض برای بسیاری از کارهای آماری دست رسانی در نظر گرفته شده است. اما برای بعضی دیگر از پردازش های آماری باید پکیج مربوطه را از اینترنت دانلود و نصب نمود. برای مثال در این قسمت مراحل کار برای دانلود و نصب پکیج ICSNP را که برای آزمونهای چند متغیره بکار می رود، بررسی می کنیم. در محیط R از مسیر زیر وارد شده،

Help\CRAN home page

و یا در مرورگر اینترنت از مسیر زیر وارد می شویم؛