

## مروری بر نرم‌افزارهای رایگان آماری

رحیل حکیمی‌مفرد<sup>۱</sup>، رضا حکیمی‌مفرد<sup>۲\*</sup>

گروه شیلات، دانشگاه آزاد اسلامی، واحد علوم و تحقیقات خوزستان، اهواز، ایران

### Reviews of Free Statistical Software

Rahil Hakimimofrad<sup>1</sup>, Reza Hakimimofrad<sup>2\*</sup>

Department of Fisheries, Khuzestan Science and Research Branch, Islamic Azad University, Ahvaz, Iran

#### Abstract

Nowadays the use of statistical software is obvious to conduct scientific research. Select statistical software according to moral considerations requires payment of fees. In this paper investigates a number of free statistical software (Such: Dap, Dataplot, Gnumeric, MicrOsiris, OpenEpi, OpenStat, PAST, PSPP, R, Salstat, SOFA, SSP, Statist, Tanagra, WebStat, WinIDAMS). Legally use the software at no charge and the results are published. The results of this study show the advantage of statistical software (R) than other free programs by their ability to function, restriction and public acceptance in use for publication of research papers.

#### Keywords

Statistics, Computer Software, Data Collection, Methodology, Data

#### چکیده

امروزه استفاده از نرم‌افزارهای آماری برای انجام تحقیقات علمی امری بدیهی به شمار می‌رود. انتخاب یک نرم‌افزار آماری با توجه به رعایت ملاحظات اخلاقی، مستلزم پرداخت هزینه است. هدف پژوهشگران در این مقاله بررسی تعدادی از نرم‌افزارهای رایگان آماری مانند: AM، Dap، Dataplot، Gnumeric، MicrOsiris، OpenEpi، OpenStat، PAST، PSPP، R، Salstat، SOFA، SSP، Statist، Tanagra، WebStat، WinIDAMS می‌باشد؛ نرم‌افزارهایی که از نظر قانونی بدون پرداخت هزینه قابل استفاده بوده و نتایج آن‌ها قابل انتشار می‌باشد. نتایج این بررسی، برتری نرم‌افزار آماری (R) را نسبت به سایر نرم‌افزارهای رایگان آماری از نظر توانایی و کارکرد، محدودیت‌ها و میزان مقبولیت عمومی در استفاده برای انتشار مقالات پژوهشی نشان می‌دهد.

#### واژگان کلیدی

آمار، نرم‌افزار، جمع‌آوری داده، روش کار، داده

#### مقدمه

آمار در پژوهش، به جمع‌آوری، تجزیه و تحلیل داده‌ها و استخراج تعمیم‌های منطقی در مورد پدیده‌ها و فرایندهای پژوهش گفته می‌شود و به‌طور وسیعی در قلمرو تمامی تحقیقات علمی به کار می‌رود، لذا امروزه لزوم استفاده از یک نرم‌افزار آماری برای انجام یک تحقیق علمی، امری بدیهی است. چنانچه از نرم‌افزارهای آماری جهت انجام مسائل آماری پژوهش استفاده شده می‌بایست مشخصات کامل آن (شامل نام، ویرایش و شرکت سازنده) قید شود. انتخاب یک نرم‌افزار آماری با توجه به رعایت ملاحظات اخلاقی در زمان انجام پژوهش و همچنین انتشار نتایج آن [۱،۲]، مستلزم پرداخت هزینه می‌باشد. در حال حاضر قیمت نرم‌افزار (SAS Analytics Pro) ۸۵۰۰ دلار، نرم‌افزار (SPSS Professional) ۱۰۳۰۰ دلار، نرم‌افزار (Minitab) ۱۳۹۵ دلار، نرم‌افزار (Stata SE) ۱۶۹۵ دلار می‌باشد [۳].

زمانی که برای یافتن یک نرم‌افزار جایگزین رایگان برای نسخه‌های تجاری جستجو می‌کنیم، با سه اصطلاح (Open Source، Freeware و Free Software) مواجه می‌شویم. نکته مهمی که باید به آن توجه شود این است که کلمه رایگان

<sup>۱</sup> دانشجوی دکتری داروسازی حرفه‌ای، دانشگاه شهید بهشتی، تهران، ایران.

<sup>۲\*</sup> نویسنده مسؤول: گروه شیلات، دانشگاه آزاد اسلامی، واحد علوم و تحقیقات خوزستان، اهواز، ایران. تمامی درخواست‌ها به نشانی "r.hakimimofrad@yahoo.com" ارسال شود.



(Free) به معنی مجانی بودن یا رایگان بودن محصول نیست بلکه به معنی آزاد بودن نرم افزار برای اجرا، تغییر، اصلاح، تکثیر و توزیع آن می باشد. در برخی موارد نرم افزارهای آزاد (به دلیل ابهام در معنی Free) به اشتباه به عنوان نرم افزارهای رایگان و احتمالاً بی ارزش تلقی می شوند، به همین دلیل امروزه این نرم افزارها به متن باز (Open Source) معروف هستند. در واقع در نرم افزارهای آزاد بحث قیمت مورد نظر نیست بلکه آزادی مطرح است. نرم افزارهای (Freeware) از نظر مالی رایگان می باشند ولی آزادی عمل نرم افزارهای (Free Software) را ندارند.

در این مقاله به بررسی اجمالی نرم افزارهای رایگان آماری پرداخته می شود؛ نرم افزارهایی که از نظر قانونی بدون پرداخت هزینه قابل استفاده بوده و نتایج آن ها قابل انتشار می باشد. تعدادی از نرم افزارهای معرفی شده، متن باز هستند لذا کدهای ساخت نرم افزار نیز برای ایجاد تغییرات دلخواه یا مورد نیاز در دسترس می باشد. نرم افزارهای رایگان جایگزین کاربردی نرم افزارهای تجاری می باشند. نتایج خروجی این دسته از نرم افزارها با نسخه های تجاری یکسان است. انتشار این دسته از نرم افزارها معمولاً توسط ارگان های دولتی، سازمان های غیردولتی (NGOs) مانند یونسکو (UNESCO)، دانشگاه ها و برخی از افراد صورت می گیرد. برخی از بسته های نرم افزاری برای اهداف خاص (به عنوان مثال، تجزیه و تحلیل سری های زمانی، تحلیل عامل، محاسبه توزیع های احتمال و غیره) توسعه یافته و برخی دیگر برای انجام طیف گسترده ای از تحلیل های آماری.

### تاریخچه نرم افزارهای رایگان آماری

برخی از نرم افزارهای رایگان توسط بخش های دولتی یا سازمان های غیردولتی طراحی و انتشار می یابند مانند نرم افزار (Epi Info) که از طرف مرکز کنترل و پیشگیری از بیماری (CDC) [۴] حمایت می شود و نرم افزار (IDAMS) که از طرف یونسکو منتشر می شود [۵]. برخی دیگر از نرم افزارها توسط شرکت های کوچک و یا دانشگاه ها انتشار می یابند مانند نرم افزار (Instat) [۶] و نرم افزار (Irristat) [۷]. دو نرم افزار دیگر به نام های (R) [۸] و (PSPP) [۹] به عنوان بخشی از پروژه گنو (GNU: GNU's Not Unix) توسط افراد زیادی که بسیاری از آن ها به صورت داوطلبانه همکاری می کنند، توسعه یافته اند. هر دو نرم افزار علاوه بر رایگان بودن، به صورت متن باز در اختیار عموم قرار گرفته اند. تعداد زیادی از نرم افزارهای آماری نیز توسط افراد مختلف به صورت شخصی انتشار یافته اند مانند Zelig، SOFA، PSPP، OpenStat، MicrOsiris و Easyreg [۱۵-۱۰]. نرم افزار IDAMS، بیشتر با هدف کمک به پژوهشگران کشورهای جهان سوم طرف یونسکو منتشر می شود. نرم افزارهای OpenStat و Instat با اهداف کمک آموزشی توسعه یافته اند [۱۲۶]. برخی از نرم افزارهای آماری رایگان برای اهداف خاصی ساخته شده اند اما به طور کلی می توان از آن ها برای سایر موارد نیز استفاده نمود مانند Irristat که برای تجزیه و تحلیل های کشاورزی ساخته شده است و یا Epi Info برای اهداف مرتبط به بهداشت عمومی و سلامت اجتماعی توسعه یافته است. برخی دیگر مانند PSPP، R و Osiris به منظور تجزیه و تحلیل های کلی در حال توسعه می باشند. از این بسته های نرم افزاری رایگان در تعدادی از مقالات منتشر شده در نشریات علمی استفاده شده است. به عنوان مثال از OpenStat [۱۶-۱۸]، Irristat [۱۹]، EasyReg [۲۰-۲۲]، نرم افزار R [۲۳-۲۵] در بسیاری از مقالات استفاده شده است که با جستجو در پایگاه های نمایه سازی، می توان مقالات زیادی را که به این نرم افزارهای رایگان ارجاع داده اند، پیدا نمود. البته برخی از نرم افزارها مانند Microsirsis به نظر نمی رسد در تحقیقات علمی تاکنون مورد استفاده قرار گرفته باشد.

در بررسی پیشینه مطالعات صورت گرفته جهت بررسی نرم افزارهای رایگان آماری می توان به دو مقاله Zhu و Kuljaca و Grant اشاره نمود [۲۶، ۲۷]. Zhu و Kuljaca برخی از ویژگی های مثبت (سهولت استفاده، تعداد روش آماری پشتیبانی شده، امکان توسعه نرم افزار و غیره) این نرم افزارها را بررسی کردند [۲۶]. البته در حال حاضر قابلیت های آماری نرم افزارهای بررسی شده در این مقاله پیشرفت قابل ملاحظه ای کرده اند. Grant نیز به بررسی نرم افزار R پرداخته است [۲۷]. در پژوهش دیگری که در آن به بررسی نرم افزارهای آماری تجاری پرداخته شده است، بسته نرم افزاری R نیز (که رایگان می باشد) مورد مطالعه قرار گرفته است [۲۸]. در بررسی دیگری، دقت نرم افزار EasyReg در مقایسه با سایر نرم افزارها، مورد بررسی قرار گرفته است [۲۹]. خروجی نرم افزارهای آماری نیز در یک بررسی مورد مقایسه قرار گرفته است [۳۰] که در تمامی نرم افزارهای مورد بررسی توانایی استفاده از خروجی نرم افزار اکسل به عنوان فایل ورودی (CSV) وجود داشت همچنین نتایج رگرسیون و همبستگی این نرم افزارها مشابه هم گزارش شد. مطالعات دیگری نیز وجود دارد که در آن به بررسی نسخه های تجاری بسته های نرم افزاری Stata، SAS، SPSS و غیره پرداخته شده است [۳۱، ۳۲].

### انتخاب یک نرم‌افزار آماری رایگان

انتخاب یک نرم‌افزار آماری رایگان تا حدود زیادی به قابلیت انجام آزمون آماری مورد نظر محقق توسط نرم‌افزار بستگی دارد اما در نسخه‌های تجاری به دلیل گستردگی و تنوع فرامین و دستورات آماری موجود در نرم‌افزار این عامل، یک عامل محدودکننده محسوب نمی‌شود. یکی دیگر از موارد مهم در انتخاب نرم‌افزار آماری، رابط کاربری نرم‌افزار می‌باشد. رابط کاربری بسیاری از این بسته‌های نرم‌افزاری گرافیکی یا گزینگان محور (Menu driven) می‌باشد ولی رابط کاربری نرم‌افزار R متنی می‌باشد و دستورات از طریق زبان برنامه‌نویسی (Code driven) به برنامه داده می‌شود. یادگیری نرم‌افزارهای دارای رابط کاربری گرافیکی در کمتر از چند ساعت امکان پذیر است ولی برای یادگیری نرم‌افزار R نیاز به صرف زمان طولانی‌تری می‌باشد. تعدادی از این نرم‌افزارها R، PSPP و Epi Info دارای بسته‌های آموزشی و راهنما استفاده می‌باشند که در یادگیری اصول اولیه برنامه به کاربر کمک می‌کند. بسیاری از آنها نیز دارای راهنمای برخط و صفحات آموزشی در فضای مجازی می‌باشند مانند نرم‌افزارهای EasyReg، R، Microsirir، Zelig، Vista، WinIdams، OpenStat و PSPP.

### محدودیت‌های استفاده از نرم‌افزارهای آماری رایگان

بسیاری از این نرم‌افزارهای آماری دارای محدودیت‌هایی می‌باشند. در نرم‌افزار (IDAMS) محدودیت طول متغیرها ۹ رقم می‌باشد [۳۳]، لذا قبل از تجزیه و تحلیل نیاز به دستکاری داده‌ها می‌باشد. در نرم‌افزارهای Instat، Epidata، Easyreg و مدیریت داده‌های گم‌شده (Missing data) امکان‌پذیر نیست [۳۰]. در نرم‌افزار EpiInfo تعداد زیادی از روش‌ها و فرامین آماری گنجانده شده است در حالی که این برنامه فاقد ارزیابی همبستگی داده‌ها بوده و نتایج آن از طریق رگرسیون قابل مشاهده است [۳۴]. در واقع این نرم‌افزار جدول همبستگی بین متغیرهای چندگانه را ایجاد نمی‌کند. جهت استفاده از نرم‌افزار Zelig، نیاز به نصب نرم‌افزار R می‌باشد زیرا این برنامه از دستورات زبان R استفاده می‌کند. البته نرم‌افزار MicroSirir نیز جهت ارائه خروجی تا حدودی به R وابسته است. در نرم‌افزارهایی که توسط افراد به صورت خصوصی انتشار یافته است یکی از محدودیت‌های خاص، پشتیبانی و پاسخگویی نامناسب سازنده می‌باشد. اگر سؤال یا مشکلی برای کاربر ایجاد شود تا رسیدن به پاسخ، ممکن است زمان زیادی صرف شود.

### معرفی نرم‌افزارهای آماری رایگان

تعداد زیادی از نرم‌افزارهای آماری به صورت رایگان عرضه می‌شوند. نرم‌افزارهای عمومی که قابلیت انجام طیف گسترده‌ای از تجزیه و تحلیل‌های آماری را در خود دارند تا نرم‌افزارهای کوچکی که برای انجام یک تحلیل خاص انتشار یافته‌اند. دسته‌ای از این نرم‌افزارها تنها برای محاسبات ساخته شده‌اند و نقش یک ماشین حساب نرم‌افزاری را دارند و دسته‌ای دیگر به صورت یک افزونه (پلاگین) بوده و قابلیت اجرا به صورت مستقل را ندارند. برخی از نرم‌افزارها نیز صرفاً برای یک گروه خاص مانند پژوهشگران علوم زیستی، علوم اجتماعی، اپیدمیولوژی، دیرنه‌شناسی و غیره توسعه یافته‌اند. در ادامه تعدادی از نرم‌افزارهای عمومی که قابلیت انجام طیف گسترده‌ای از تجزیه و تحلیل‌های آماری را دارند به شکل مختصر معرفی می‌شوند. در این فهرست نرم‌افزارهای قدیمی و نرم‌افزارهایی که از تاریخ انتشار آخرین نسخه آن‌ها بیش از ۳ سال گذشته است، نرم‌افزارهایی که برای تولید محتوا به زبان فارسی کاربردی ندارد، معرفی نشده است.

AM: از این نرم‌افزار آماری بیشتر برای ارزیابی حجم انبوهی از داده‌ها که به صورت نظرسنجی جمع‌آوری شده‌اند، استفاده می‌شود.

Dap: نسخه رایگان نرم‌افزار تجاری SAS می‌باشد. بسیاری از دستورات SAS در Dap قابل اجرا می‌باشد. این نرم‌افزار صرفاً برای سیستم‌های عامل لینوکس و یونیکس طراحی شده است.

Dataplot: این نرم‌افزار دارای قابلیت‌های فراوان آماری و گرافیکی برای تجزیه و تحلیل‌های ساده و مدل‌سازی غیرخطی می‌باشد.

Gnumeric: یک نرم‌افزار آماری صفحه گسترده با ویژگی‌های آماری به مراتب بهتر از Excel. ۶۰ تابع آماری اضافه نسبت به اکسل در این نرم‌افزار موجب شده است تا محاسبات مالی، مخابراتی، تجزیه و تحلیل‌های پیشرفته، ابزار شبیه‌سازی مونت‌کارلو و غیره، به راحتی قابل انجام باشد.



**MicroSiris:** یک بسته آماری کم حجم و جامع که برای کاربران سیستم عامل ویندوز انتشار یافته است. مهم ترین ویژگی این نرم افزار، ورود داده ها از سایر برنامه ها مانند SAS، SPSS، Excel، OSIRIS، IDAMS می باشد. البته خروجی این نرم افزار نیز قابل خواندن برای سایر نرم افزارها می باشد. دستورات و فرامین داده کاوی برای تجزیه و تحلیل بازار به خوبی در این نرم افزار قرار داده شده است.

**OpenEpi:** نرم افزار رایگان و متن باز که بیشتر برای اهداف پزشکی توسعه و انتشار یافته است. این نرم افزار مستقل از سیستم عامل قابل اجرا است.

**OpenStat:** نرم افزاری برای انجام کارهای عمومی آماری که بر روی نسخه های مختلف ویندوز ۷، ۸ و Xp، و همچنین لینوکس قابل اجرا می باشد. رابط کاربری آن شبیه SPSS بوده و دارای فایل راهنمای جامعی جهت کار با نرم افزار است.

**PAST:** یک نرم افزار آماری با کاربری ساده که برای اهداف دیرینه شناسی انتشار یافته است. فرامین رایج آماری، رسم و مدل سازی توابع، انواع مختلف آزمون های آماری، جداول احتمال، آمار چند متغیره، تجزیه و تحلیل سری های زمانی، تجزیه و تحلیل هندسی و غیره بخشی از توانایی های این نرم افزار می باشد.

**PSPP:** جایگزین رایگان نرم افزار تجاری SPSS می باشد با این تفاوت که تنها بخش کوچکی از دستورات آماری SPSS را در خود دارد. رابط کاربری آن شباهت زیادی به SPSS دارد و در نگاه اول ممکن است کاربران را دچار اشتباه کند.

**R:** یک زبان و محیط نرم افزاری برای محاسبات آماری و گرافیکی می باشد که توسط تلاش گروهی جمعی از محققین و علاقمندان علوم محاسباتی در حال توسعه می باشد. این نرم افزار قابلیت نصب و اجرا بر روی سیستم عامل های ویندوز، لینوکس، مک و غیره را دارد. برای آموزش این نرم افزار فایل های نوشتاری و تصویری زیادی در شبکه اینترنت وجود دارد. بیش از ۹۰ پایگاه مجزا در سراسر دنیا برای آموزش، پاسخ گویی، رفع اشکال و دانلود این نرم افزار وجود دارد. سرور دانلود این نرم افزار در ایران پایگاه اینترنتی دانشگاه فردوسی مشهد می باشد.

**Salstat:** قابلیت نصب و اجرا بر روی سیستم عامل های مختلف را دارد و از این نظر محدودیتی برای این برنامه وجود ندارد. رابط کاربری ساده ای دارد لذا کار با نرم افزار راحت و بدون پیچیدگی می باشد. علاوه بر انجام تجزیه و تحلیل های آماری، ترسیم انواع مختلف نمودارها نیز با این برنامه امکان پذیر است.

**SOFA:** نرم افزاری متن باز که با شعار آمار برای همه، انتشار یافته است. بر روی ویندوز، مک و لینوکس قابل نصب و اجرا می باشد. یادگیری آن بسیار ساده بوده به طوری که مهم ترین برتری این نرم افزار به حساب می آید. علاوه بر این خروجی های این نرم افزار نیز قابل تحسین است.

**SSP:** یک نرم افزار آماری ساده و کاربر پسند برای سیستم عامل های ویندوز و مک. انجام محاسبات ساده، ارزیابی توزیع احتمال، آزمون های معنی دار بودن داده ها، رگرسیون های ساده و چند متغیره و غیره بخشی از قابلیت های این نرم افزار است.

**Statist:** دارای مجموعه ای از دستورات و فرامین آماری. خروجی های این برنامه، گرافیکی و مبتنی بر متن است. تا حدودی مستقل از سیستم عامل و تحت اینترنت قابل اجرا می باشد.

**Tanagra:** بسته آماری رایگان و متن باز که بر داده کاوی متمرکز می باشد.

**WebStat:** یک نرم افزار آماری ساده مبتنی بر جاوا که برای استفاده نیاز به مرورگر می باشد. امکان دانلود و اجرای ناپرخط (Offline) نرم افزار در نسخه جدید فراهم شده است.

**WinIDAMS:** این بسته نرم افزاری توسط یونسکو برای پردازش اطلاعات عددی و تجزیه و تحلیل آماری ساده انتشار می یابد. دستکاری داده ها، روش های آماری کلاسیک و پیشرفته، جداول تعاملی، تکنیک های آماری چند متغیره و غیره بخشی از قابلیت های این نرم افزار است.

### مقایسه توانایی و کارایی نرم افزارهای آماری رایگان

جهت مقایسه نرم افزارهای آماری رایگان، پس از دانلود از پایگاه های منتشرکننده، هر یک از نرم افزارها نصب شده سپس با استفاده از راهنمای موجود در نرم افزار و همچنین داده های نمونه، کارایی و توانایی آن ها در فرامین مختلف مورد مقایسه قرار گرفت. نتایج مقایسه توانایی هر یک از نرم افزارها برای روش های مختلف تجزیه و تحلیل واریانس (ANOVA) در جدول ۱، پشتیبانی هر یک از نرم افزارها از روش های مختلف تجزیه و تحلیل رگرسیون (Regression) در جدول ۲، توانایی انجام تجزیه و تحلیل مختلف

سری‌های زمانی (Time Series Analysis) در جدول ۳، توانایی هر یک از نرم‌افزارها در ترسیم انواع مختلف نمودارها در جدول ۴ و قابلیت هر یک از نرم‌افزارها در پشتیبانی از آزمون‌های مختلف آماری در جدول ۵ آمده است. در جدول ۶ نشانی دریافت نرم‌افزارهای بررسی شده آمده است. برای تمامی نرم‌افزارها آخرین نسخه ارائه شده توسط پایگاه منتشر کننده تا پایان سال ۲۰۱۳ میلادی، جهت مقایسه در نظر گرفت شد. در این بررسی نرم‌افزار R با توانایی انجام ۳۸ دستور مختلف بالاتر از سایر نرم‌افزارهای رایگان قرار گرفت.

جدول ۱: پشتیبانی از روش‌های مختلف تجزیه و تحلیل آنالیز واریانس (ANOVA)

| تجزیه و تحلیل (ANOVA)* |          |             |     |        |         |         | نام نرم‌افزار |
|------------------------|----------|-------------|-----|--------|---------|---------|---------------|
| Latin squares          | Post-hoc | Mixed model | GLM | MANOVA | Two-way | One-way |               |
| خیر                    | خیر      | خیر         | خیر | خیر    | بلی     | بلی     | ADaMSoft      |
| خیر                    | خیر      | خیر         | خیر | خیر    | بلی     | بلی     | Epi Info      |
| بلی                    | بلی      | نامشخص      | بلی | بلی    | بلی     | بلی     | PSPP          |
| بلی                    | بلی      | بلی         | بلی | بلی    | بلی     | بلی     | R             |
| نامشخص                 | بلی      | بلی         | خیر | بلی    | بلی     | بلی     | Sage          |
| خیر                    | خیر      | نامشخص      | خیر | خیر    | خیر     | بلی     | Salstat       |
| خیر                    | خیر      | خیر         | خیر | خیر    | خیر     | بلی     | SciPy         |
| بلی                    | بلی      | نامشخص      | خیر | خیر    | بلی     | بلی     | SOCR          |
| خیر                    | خیر      | نامشخص      | خیر | خیر    | خیر     | بلی     | SOFA          |

\*: آنالیز واریانس یک طرفه (One-way)، آنالیز واریانس دو طرفه (Two-way)، آنالیز واریانس چندمتغیره (MANOVA)، مدل خطی عمومی (GLM)، نمونه‌های آمیخته (Mixed model)، آزمون‌های تعقیبی (Post-hoc)، مربع‌های لاتین (Latin squares).

در این میان نرم‌افزار R توانایی انجام تمامی آزمون‌های تحلیل واریانس را دارد. در بین نرم‌افزارهای تجاری تنها نرم‌افزار SAS توانایی انجام چنین طیف گسترده‌ای از آزمون‌های تحلیل واریانس را دارد.

جدول ۲: پشتیبانی از روش‌های مختلف تجزیه و تحلیل رگرسیون (Regression)

| تجزیه و تحلیل (Regression)* |         |     |        |          |          |                  |                  |          |                  |                   |                  |                  | نام نرم‌افزار |
|-----------------------------|---------|-----|--------|----------|----------|------------------|------------------|----------|------------------|-------------------|------------------|------------------|---------------|
| MLR <sup>7</sup>            | Poisson | Cox | Probit | Quantile | Stepwise | LAD <sup>6</sup> | GLM <sup>5</sup> | Logistic | NLS <sup>4</sup> | 2SLS <sup>3</sup> | WLS <sup>2</sup> | OLS <sup>1</sup> |               |
| خیر                         | خیر     | خیر | خیر    | خیر      | بله      | خیر              | خیر              | بلی      | بلی              | خیر               | بلی              | بلی              | ADaMSoft      |
| خیر                         | خیر     | خیر | خیر    | خیر      | خیر      | خیر              | خیر              | بلی      | خیر              | خیر               | خیر              | بلی              | Epi Info      |
| بلی                         | خیر     | خیر | بلی    | بلی      | بلی      | بلی              | خیر              | بلی      | بلی              | بلی               | بلی              | بلی              | Gretl         |
| خیر                         | خیر     | خیر | خیر    | خیر      | خیر      | خیر              | خیر              | خیر      | خیر              | خیر               | خیر              | بلی              | PSPP          |
| بلی                         | بلی     | بلی | بلی    | بلی      | بلی      | بلی              | بلی              | بلی      | بلی              | بلی               | بلی              | بلی              | R             |
| خیر                         | بلی     | خیر | بلی    | بلی      | بلی      | بلی              | بلی              | بلی      | بلی              | بلی               | بلی              | بلی              | Sage          |
| خیر                         | خیر     | خیر | خیر    | خیر      | خیر      | خیر              | خیر              | خیر      | خیر              | خیر               | خیر              | خیر              | Salstat       |
| خیر                         | خیر     | خیر | خیر    | خیر      | بلی      | خیر              | خیر              | خیر      | بلی              | خیر               | خیر              | بلی              | SciPy         |
| خیر                         | خیر     | خیر | خیر    | خیر      | خیر      | خیر              | خیر              | بلی      | خیر              | خیر               | خیر              | بلی              | SOCR          |

\*: در این مقایسه نرم‌افزار R در انجام مدل‌های مختلف رگرسیون، از سایر نرم‌افزارها متمایز است. در بین نرم‌افزارهای تجاری تنها نرم‌افزار SAS، قابل مقایسه با R می‌باشد.

- |   |                             |                               |
|---|-----------------------------|-------------------------------|
| 1. Ordinary least squares                 | 4. Non-linear least squares | 7. Multiple linear regression |
| 2. Weighted least squares                 | 5. Generalized linear model |                               |
| 3. Two stage least squares (2SLS or TSLS) | 6. Least absolute deviation |                               |



جدول ۳: پشتیبانی از روش‌های مختلف تجزیه و تحلیل سری‌های زمانی (Time series analysis)

| تجزیه و تحلیل سری‌های زمانی* |     |                    |                |       |       | نام نرم‌افزار |
|------------------------------|-----|--------------------|----------------|-------|-------|---------------|
| Multivariate GARCH           | VAR | Cointegration test | Unit root test | GARCH | ARIMA |               |
| نامشخص                       | بلی | بلی                | بلی            | بلی   | بلی   | gretl         |
| بلی                          | بلی | بلی                | بلی            | بلی   | بلی   | R             |
| بلی                          | بلی | بلی                | بلی            | بلی   | بلی   | Sage          |
| خیر                          | خیر | خیر                | خیر            | خیر   | خیر   | Salstat       |
| خیر                          | خیر | خیر                | نامشخص         | خیر   | خیر   | SciPy         |
| نامشخص                       | خیر | خیر                | خیر            | خیر   | خیر   | SOCR          |

\*: نرم‌افزارهای R در این زمینه نیز توانایی انجام تمامی تجزیه و تحلیل‌های مربوط به سری‌های زمانی را دارد. در بین نسخه‌های تجاری، نرم‌افزار SPSS در این زمینه محدود است. نرم‌افزار SAS هم تراز با R بوده و پرکاربردترین نرم‌افزار برای انجام تحلیل‌ها مربوط به سری‌های زمانی Minitab می‌باشد.

جدول ۴: پشتیبانی از روش‌های مختلف ترسیم نمودارهای آماری

| نمودارهای آماری |            |           |             |          |           | نام نرم‌افزار |
|-----------------|------------|-----------|-------------|----------|-----------|---------------|
| Scatterplot     | Line chart | Histogram | Correlogram | Box plot | Bar chart |               |
| بلی             | بلی        | بلی       | بلی         | بلی      | بلی       | ADaMSOft      |
| بلی             | بلی        | بلی       | خیر         | خیر      | بلی       | Epi Info      |
| بلی             | بلی        | بلی       | بلی         | بلی      | بلی       | gretl         |
| بلی             | بلی        | بلی       | بلی         | بلی      | بلی       | R             |
| بلی             | بلی        | بلی       | بلی         | بلی      | بلی       | Sage          |
| بلی             | بلی        | بلی       | خیر         | بلی      | بلی       | SciPy         |
| بلی             | بلی        | بلی       | بلی         | بلی      | بلی       | SOCR          |

جدول ۵: پشتیبانی از آزمون‌های مختلف جهت تجزیه و تحلیل

| SOFA                | SOCR | SciPy            | R   | PSPP   | OpenEpi          | Epi Info         | ADaMSOft | نوع آزمون       |
|---------------------|------|------------------|-----|--------|------------------|------------------|----------|-----------------|
| نامشخص <sup>۱</sup> | بلی  | بلی              | بلی | بلی    | بلی <sup>۱</sup> | بلی <sup>۱</sup> | بلی      | آمار توصیفی     |
| بلی                 | بلی  | بلی <sup>۳</sup> | بلی | نامشخص | بلی <sup>۲</sup> | بلی              | بلی      | آمار ناپارامتری |
| خیر                 | خیر  | خیر              | بلی | نامشخص | خیر              | خیر              | بلی      | کنترل کیفیت     |
| خیر                 | بلی  | خیر              | بلی | نامشخص | خیر              | بلی              | بلی      | تحلیل بقا       |
| خیر                 | بلی  | خیر              | بلی | نامشخص | خیر              | خیر              | بلی      | تحلیل خوشه‌ای   |
| خیر                 | بلی  | بلی              | بلی | نامشخص | خیر              | بلی              | بلی      | داده‌پردازی     |

<sup>۱</sup>: فاقد آزمون نرمال بودن داده‌ها، <sup>۲</sup>: فقط تجزیه و تحلیل جداول احتمال، <sup>۳</sup>: فاقد امکان تجزیه و تحلیل جداول احتمال

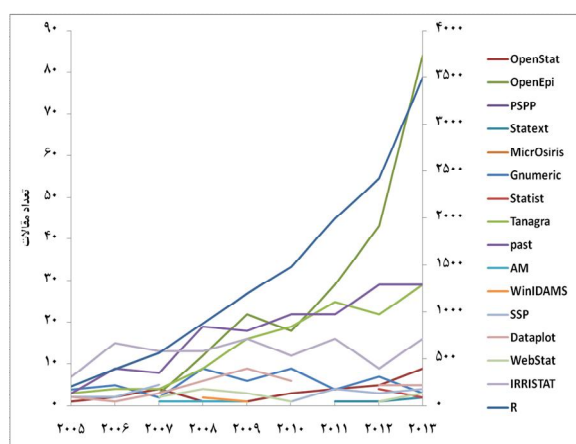
جدول ۶: نشانی دریافت نرم‌افزارها

| نشانی دریافت (قابل دسترسی در تاریخ: ۱۳۹۲/۱۰/۱)  | نام نرم‌افزار |
|---|---------------|
| <a href="http://adamsoft.sourceforge.net">http://adamsoft.sourceforge.net</a>                                   | ADaMSOft      |
| <a href="http://www.itl.nist.gov/div898/software/dataplot">http://www.itl.nist.gov/div898/software/dataplot</a> | Dataplot      |
| <a href="http://wwwn.cdc.gov/epiinfo">http://wwwn.cdc.gov/epiinfo</a>   | Epi Info      |
| <a href="http://gretl.sourceforge.net">http://gretl.sourceforge.net</a>   | Gretl         |
| <a href="http://www.openepi.com/v37/Menu/OE_Menu.htm">http://www.openepi.com/v37/Menu/OE_Menu.htm</a>           | OpenEpi       |
| <a href="https://www.gnu.org/software/pspp">https://www.gnu.org/software/pspp</a>                               | PSPP          |
| <a href="http://www.r-project.org">http://www.r-project.org</a>   | R             |
| <a href="http://www.sagemath.org">http://www.sagemath.org</a>   | Sage          |
| <a href="http://www.salstat.com">http://www.salstat.com</a>   | Salstat       |
| <a href="http://jwork.org/scavis">http://jwork.org/scavis</a>   | SCaViS        |
| <a href="http://www.scipy.org">http://www.scipy.org</a>   | SciPy         |
| <a href="http://www.socr.ucla.edu">http://www.socr.ucla.edu</a>   | SOCR          |
| <a href="http://www.sofastatistics.com/home.php">http://www.sofastatistics.com/home.php</a>                     | SOFA          |



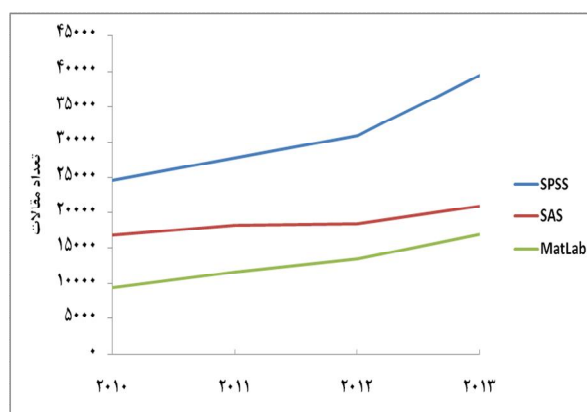
### استفاده از نرم‌افزارهای آماری رایگان

جهت بررسی آمار استفاده از نرم‌افزارهای آماری رایگان، تعداد مقالات منتشر شده در پایگاه sciencedirect که از هر یک از نرم‌افزارهای آماری رایگان استفاده شده است، مورد بررسی قرار گرفت. تصویر ۱، تعداد مقالات انتشار یافته بین سال‌های ۲۰۰۵ تا ۲۰۱۳ میلادی را نشان می‌دهد که از نرم‌افزارهای آماری رایگان جهت تجزیه و تحلیل داده‌ها استفاده شده است. در سال ۲۰۱۳، ۹۵٪ پژوهش‌های منتشر شده توسط نرم‌افزارهای آماری رایگان با استفاده از نرم‌افزار R، ۲٪ با استفاده از OpenEpi و ۳٪ با استفاده از سایر نرم‌افزارهای آماری رایگان انجام شده است. در تصویر ۲، آمار استفاده از نرم‌افزارهای تجاری در انتشار مقالات علمی جهت مقایسه با تصویر ۱، آمده است. اگر چه استفاده از R نسبت به ۷ سال گذشته بیش از ۱۶ برابر رشد داشته است ولی در حال حاضر قابل مقایسه با نرم‌افزار SPSS نیست، در حالی که توانایی نرم‌افزار R قابل رقابت با نسخه‌های تجاری نرم‌افزارهای مشابه می‌باشد. استفاده از نرم‌افزار R طی ۴ سال گذشته ۱/۷ برابر و نرم‌افزار SPSS، ۱/۴ برابر رشد داشته است که می‌تواند نشان‌دهنده تغییر سلیقه کاربران در آینده باشد.



تصویر ۱: آمار استفاده از نرم‌افزارهای آماری رایگان در مقالات پایگاه Science Direct

محور عمودی سمت چپ، تعداد مقالاتی را نشان می‌دهد که از نرم‌افزار (R) استفاده شده است.



تصویر ۲: آمار استفاده از نرم‌افزارهای تجاری در مقالات پایگاه Science Direct

### نتیجه‌گیری

اگر تاکنون در انجام پژوهش‌های خود از بسته‌های نرم‌افزاری تجاری بدون رعایت قوانین حق مؤلف (Copy Right) استفاده نموده‌اید (اگرچه به صورت ناآگاهانه بوده است) و تصمیم به استفاده از بسته‌های نرم‌افزاری رایگان دارید در این صورت R یک نرم‌افزار رایگان برای محاسبات آماری و تحلیل داده‌ها با محیط برنامه‌نویسی می‌باشد و با توجه به مقایسه صورت گرفته قوی‌ترین

نرم افزار آماری رایگان می باشد. از مهم ترین ویژگی های این نرم افزار می توان به متن باز بودن آن اشاره کرد که به کاربران اجازه می دهد با استفاده از کدهای زبان "C" آن را تغییر دهند و به نتیجه مطلوب خود دست یابند. اما از معایب این نرم افزار می توان به محیط کدنویسی آن اشاره کرد، که موجب می شود کاربران عادی نتوانند از آن استفاده کنند و نیاز به آموزش داشته باشند. ابزارهای مختلفی جهت کاربرپسند نمودن محیط نرم افزار R تاکنون ارائه شده است. اصطلاحاً به هر یک این ابزارها رابط کاربری گرافیکی (R GUI) گفته می شود. از این میان می توان به RStudio اشاره نمود [۳۵] که یک رابط کاربری گرافیکی را با ابزارهای قدرتمند جهت برنامه نویسی می باشد. هدف توسعه دهندگان RStudio کمک به کاربران R برای استفاده بیشتر می باشد.

### سپاسگزاری

نویسندگان مقاله مراتب تقدیر و تشکر خود را از سرکار خانم دکتر مریم محمدی (معاونت غذا و دارو دانشگاه علوم پزشکی جندی شاپور اهواز) ابراز می دارند.

### References:

1. Enjoo A. Assessing Intellectual Property, Publication Ethics, and Related Legal provisions in the Islamic Republic of Iran. MEDIA. 2011;2(2):36-44. [In Persian]
2. Abbasi M. Ethical Codes for Medical Research. Medical Ethics Series. Tehran: Hoghooghi; 2006. P. 5-8. [In Persian]
3. ProStatServices. A Review of the Top Five Statistical Software Systems [Internet]. 2012 [Cited 2012 Nov 24]. Available from: <http://www.prostatservices.com/statistical-consulting/articles-of-interest/a-review-of-the-top-five-statistical-software-systems>.
4. Epi Info. CDC [Internet]. 2008 [Cited 2013 Nov 21]. Available from: <http://www.cdc.gov/epiinfo/index.htm>
5. IDAMS Statistical Software [Internet]. 2008 [Cited 2013 Nov 21]. Available from: [http://portal.unesco.org/ci/en/ev.php-URL\\_ID=5416&URL\\_DO=DO\\_TOPIC&URL\\_SECTION=201.html](http://portal.unesco.org/ci/en/ev.php-URL_ID=5416&URL_DO=DO_TOPIC&URL_SECTION=201.html)
6. InStat - An Interactive Statistical Package. Statistical Services Centre - University of Reading [Internet]. 2009 [Cited 2014 Dec 2]. Available from: <http://www.ssc.rdg.ac.uk/software/instat/instat.html>.
7. Irristat. International Rice Research Institute, Biometrics and Bioinformatics Unit [Internet]. 2014 [Cited 2014 Dec 2]. Available from: <http://www.irri.org>.
8. Rossiter D. Introduction to the R Project for Statistical Computing for Use at the ITC [Internet]. 2012 [Cited 2014 Dec 1]. Available from: <http://cran.r-project.org/doc/contrib/Rossiter-RIntro-ITC.pdf>.
9. Pfaff B. PSPP [Internet]. 2013 [Cited 2014 Nov 21]. Available from: <https://www.gnu.org/software/pspp>.
10. Easy Reg International. Herman Bierens, Penn State University [Internet]. 2008 [Cited 2014 Dec 2]. Available from: <http://econ.la.psu.edu/~hbierens/EASYREG.HTM>
11. Van Eck N. MicOsiris. Van Eck Computer Consulting [Internet]. 2013 [Cited 2014 Dec 2]. Available from: <http://www.microsiris.com>.
12. Miller B. OpenStat [Internet]. 2009 [Cited 2013 Nov 24]. Available from: <http://www.statpages.org/miller/openstat>.
13. PSPP. 2014. Available from: <http://www.gnu.org/software/pspp>.
14. Paton-Simpson G. SOFA- Statistics Open for All [Internet]. 2014 [Cited 2014 Dec 2]. Available from: <http://www.sofastatistics.com/home.php>
15. Kosuke I, King G, Lau O. Zelig: Everyone's Statistical Software [Internet]. 2006 [Cited 2014 Dec 2]. Available from: <http://gking.harvard.edu/zelig>.
16. Mark E. Future Salary and US Residency Fill Rate Revisited. JAMA. 2008;300(10):1131-1132. doi:10.1001/jama.300.10.1131. PMID: 18780840.
17. Bielaszewska M, Sinha B, Kuczius T, Karch H. Cytolethal Distending Toxin from Shiga Toxin-Producing Escherichia coli O157 Causes Irreversible G2/M Arrest, Inhibition of Proliferation and Death of Human Endothelial Cells. Infection and Immunity. 2005 Jan [Cited 2013 Nov 21];73(1):552-562. Available from: <http://iai.asm.org/cgi/content/full/73/1/552>.
18. Toscano CD, Kingsley PJ, Marnett LJ, Bosetti F. NMDA-induced Seizure Intensity is Enhanced in COX-2 Deficient Mice. Neurotoxicology. 2008 [Cited 2014 Dec 2];29(6):1114-1120. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2587528>



19. FAO Plant Production and Protection Paper No. 174, Rome [Internet]. 2003. [Cited 2014 Dec 2]. Available from: <http://www.fao.org/DOCREP/005/Y4391E/y4391>.
20. Gambardella A, Hall BH. Proprietary versus Public Domain Licensing of Software and Research Products. *Research Policy*. 2006 [Cited 2014 Dec 4];35(6):875-892. Available from: <http://repositories.cdlib.org/postprints/1865>.
21. Wen-Chi L, Chang T. Rational Bubbles in the Korea Stock Market? Further Evidence based on Nonlinear and Nonparametric Cointegration Tests. *Economics Bulletin* [Internet]. 2008 [Cited 2014 Dec 4];3(34):1-12. Available from: <http://economicsbulletin.vanderbilt.edu/2008/volume3/EB-08C30021A.pdf>
22. Itoa H, Lee D. Assessing the Impact of the September 11 Terrorist Attacks on U.S. Airline Demand. *Journal of Economics and Business*. 2005 [Cited 2014 Dec 4];57(1):75-95. Available from: <http://dx.doi.org/10.1016/j.jeconbus.2004.06.003>.
23. Handcock MS, Hunter DR, Butts CT, Goodreau SM, Morris M. statnet: Software Tools for the Representation, Visualization, Analysis and Simulation of Network Data. *J Stat Softw*. 2008 [Cited 2014 Dec 4];24(1):1548-7660. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2447931>.
24. Hume ME, Scanlan CM, Harvey RB, Andrews K, Snodgrass JD, Nalian AG, Martynova-Van Kley A, Nisbet DJ. Denaturing Gradient Gel Electrophoresis as a Tool to Determine Batch Similarity of Probiotic Cultures of Porcine Cecal Bacteria. *Applied and Environmental Microbiology*, 2008 Aug [Cited 2014 Dec 4]; 74(16):5241-5243. Available from: <http://aem.asm.org/cgi/content/abstract/74/16/5241>.
25. Bylesjö M, Nicholson JK, Holmes E, Trygg J. K-OPLS package: Kernel-based orthogonal projections to latent structures for prediction and interpretation in feature space. *Journal of BMC Bioinformatics*. 2008;9(1):106.
26. Zhu X, Kuljaca O. A Short Preview of Free Statistical Software Packages for Teaching Statistics to Industrial Technology Majors. *Journal of Industrial Technology*. 2005;21(2):21-22.
27. Grant F. Free Statistics Software, Yours, Free to keep. *Scientific Computing World* [Internet]. 2004 [Cited 2014 Dec 4]. Available from: [http://www.scientific-computing.com/scwsepoct04free\\_statistics.htm](http://www.scientific-computing.com/scwsepoct04free_statistics.htm).
28. Wegman EJ, Solka JL. Statistical Software for Today and Tomorrow [Internet]. 2005 [Cited 2014 Dec 1]. Available from: <http://www.galaxy.gmu.edu/> listed as "A Guide to Statistical Software".
29. Choia HS, Kiefer NM. Software Evaluation: EasyReg International. *International Journal of Forecasting*. 2005;21(3):609-616.
30. Gene S. Comparing Free Statistical Software for Data Sets with no Missing Values and Comparing Free Statistical Software, Handling Missing Data. Both Available Here Free Software [Internet]. 2006 [Cited 2014 Dec 4]. Available from: <http://gsociology.icaap.org/methods/soft.html>.
31. Acock AC. SAS, Stata, SPSS: A Comparison. *Journal of Marriage and Family*. 2005;4(67):1093-1095.
32. Wass JA. Comparative Statistical Software Review. Tabulations and Musings from your Editor's Biased Perspective [Internet]. 2006 [Cited 2014 Dec 4]. Available from: <http://www.scientificcomputing.com/comparative-statistical-software.aspx>.
33. Unesco. WinIDAMS 1.3 Reference Manual - Table of Contents [Internet]. 2008 [Cited 2014 Dec 1]. Available from: <http://www.unesco.org/webworld/portal/idams/html/english/TOC.htm>.
34. CDC. Epi Info Training Session. Using Epi Info in an Outbreak Investigation. *Advanced Analysis and Mapping* [Internet]. 2008 [Cited 2014 Dec 4]. Available from: <http://www.cdc.gov/epiinfo>.
35. RStudio, Inc [Internet]. 2013 [Cited 2013 Nov 21]. Available from: <http://www.rstudio.com>.