

تخمین تورش Bias و قابلیت تکرار با استفاده از نمونه های بیماران Spreadsheet B, version 3

خلاصه

این برنامه طراحی اولیه ای است برای تخمین تورش بین دو روش Bias. در این برنامه میتوان، از اندازه گیری های تکی و یا دوتائی (single or duplicate measurements) استفاده نمود.

همچنین در انتها میتوان قابلیت تکرار روشها را نیز تخمین زد. نمودار توزیع و نمودار نقطه ای ای تفاوت ها نیز تهیه شده است

A scatter gram and relative and absolute difference graphs.

سوابق

اندازه گیری تورش یک متد در مقابل روش قبلی، با استفاده از مقایسه نمونه های بیماران انجام میشود. این مقایسه هیچ گونه اطلاعاتی درباره اینکه این روش در مقابل یک روش مرجع Reference Method دارای تورش میباشد یا خیر بدست نمی دهد. با استفاده از اندازه گیری دوتائی نمونه ای بیماران در طیف غلظتی ای که محدوده اندازه گیری روشها را بپوشاند میتوان تخمین خوبی از تورش و قابلیت تکرار روشها بدست آورد.

طراحی برنامه

۲۰ نمونه بیمار با غلظتهائی که محدوده اندازه گیری ویا محدوده های رایج کلینیکی را بپوشاند انتخاب کنید. نمونه ها ئی با موارد شناخته شده تداخل، مانند نمونه های لیپمیک، همولیز ویا یرقانی نباید استفاده شود.

هر نمونه باید به صورت دوتائی توسط هر دو روش مقایسه ای (مرجع) و روش مورد ارزیابی، اندازه گیری شود. در این برنامه از خواننده های تکی نیز میتوان استفاده نمود ولی بدیهی است در این صورت نتیجه بدست آمده از اطمینان کمتری برخوردار بوده و قابلیت تکرار نیز قابل تخمین نخواهد بود.

نتایج را در خانه های مربوط در صفحه وارد نمائید اگر از نمونه هارا بصورت تک خواننده اندازه گیری کرده اید برای روش تست خواننده هارا در اولین ستون (E) و برای روش مقایسه ای در اولین ستون (B) وارد نمائید.

توجه: نام و مشخصات سیستم ها و کمیت مورد اندازه گیری میباید قبل از شروع محاسبات در خانه های B12-E12 و D10 وارد شود. به علاوه حداقل و حداکثر محدوده اندازه گیری متد مقایسه ای نیز باید برای نمودار وارد شود. مقادیر از پیش تعیین شده مانند ۰-۱۰۰ ممکن است کاملاً نامربوط بوده و اطلاعات نمودار را حذف نماید.

درجه بندی محورهای نمودارها بصورت خودکار تنظیم میشود ولی هرگاه تعداد نتایج بسیار زیاد باشد افزایش میزان حداقلی محورهای نمودار مهم است. این عمل از طریق روش استاندارد اکسل انجام میشود. درجه بندی محورهای X و Y نمودار باید یکسان باشد.

محاسبات

همیشه نمودارها را جهت مشاهده خواننده های دور افتاده **outliers** احتمالی بررسی نمائید

ابتدا میباید درباره اینکه محاسبه با تک خواننده ها انجام شود یا با خواننده های دوتائی تصمیم گیری نمود. (انتخاب در خانه های $K3$ و $K4$)

در صورت انتخاب محاسبه تک خواننده ، فقط خواننده های ستون اول مورد استفاده قرار میگیرند و خواننده های ستونهای C و F منظور نمیشوند.

....

برنامه پارامترهای مهم آماری را محاسبه میکند. (شکل ۱). "میزان رد کاذب" نتایج (α) **false rejection rate** میتواند بادر نظر گرفتن سطح اطمینان 99% و 95% انتخاب شود: 1% یا $5\% = \alpha$.

جهت تعیین اینکه آیا اختلاف معنی داری بین نتایج دو روش وجود دارد یا خیر ، آزمون **Student s dependent t-test** اضافه شده است. این تست بر مبنای اختلاف مشاهده شده بین جفت خواننده هادر اولین ستون هر گروه میباشد هرگاه از مدل ارزیابی اندازه گیری تکی استفاده شده باشد. جمع بندی ارزیابی با استفاده از واژه های "اختلاف معنی دار **Significant difference**" و "بدون اختلاف معنی دار **No significant difference**" ارائه میگردد.

برای معتبر بودن **Student's tdep** ، لازم است اختلاف ها ، دارای توزیع نرمال باشند. و این موضوع مستقل از توزیع نتایج و یا مشاهدات، میباشد. علاوه بر **Student s dependent t-test** ، آزمون **Wilcoxon sign rank** نیز انجام شده و مقدار p -value اندازه گیری و گزارش شده است.

تعداد خواننده های در نظر گرفته نشده **missing data** نیز بطور جداگانه هم برای روش مورد بررسی و هم روش مورد مقایسه گزارش شده است.

میانگین های روشهای مورد بررسی و مقایسه ای محاسبه شده اند. میانه و یا انحراف معیار توزیع نتایج میتوانند مورد انتخاب قرار گیرند. (M11)

coverage factor ضریب هم پوشانی و صدک بالائی **upper percentile** قابل تعیین میباشد و صدک پائینی **lower percentile** از تفرق تفاوت صدک بالائی و 100% تعیین میگردد. مقادیر پیشنهاد شده ضریب هم پوشانی و صدک بالائی عبارتند از: **upper percentile 97.5** و $k=2$

میانگین تورش ، میانگین اختلافهای نتایج هر نمونه از دوروش مقایسه ای و مورد بررسی است به جای آنکه اختلاف نتایج میانگین هریک از روشها مورد محاسبه قرار گیرد. تفاوت بین این دوروش تخمین، فقط زمانی قابل توجه میباشد که برخی از مشاهدات و نتایج در نظر گرفته نشده باشند. some observations missing. درصد تورش مربوط برای هر نمونه در ارتباط با نتایج روش مقایسه ای و میانگین هریک از مشاهدات ، محاسبه شده است.

تورش ، به عنوان تفاوت بین متد مورد بررسی و متد مقایسه ای محاسبه شده است *i.e. Test – Comparative*

رگرسیون خطی معمولی ordinary مورد استفاده قرار گرفته است. رگرسیون میتواند با استفاده از میانگین مشاهدات، و یا از اولین نتیجه از اندازه گیری های دوتائی duplicate measurements محاسبه شود. این انتخاب از طریق کلید های Y ویا N در خانه های K^3 و K^4 انجام میشود.

اگر اندازه گیری ها دوتائی duplicate measurements انجام شده اند، علاوه بر مقدار تورش ، مقدار عدم دقت هر روش نیز قابل محاسبه است. این محاسبه بر مبنای تفاوت نتایج خوانده های دوتائی انجام میشود.

رگرسیون اورتوگونال دمینگ (The orthogonal regression (Deming) امکان وجود تغییراتی در نتایج در هر دو متغیر وابسته dependent (Y) و غیر وابسته independent (X) را فراهم میسازد ولی نیازمند مشخص نمودن مقدار نسبت λ ratio $(SD_{test}/SD_{comparative})^2$ میباشد. هرگاه اندازه گیری ها دوتائی انجام شده باشند، λ از این خوانده ها محاسبه میشود. و اگر این مقدار بزرگ باشد، رگرسیون دمینگ مورد استفاده قرار میگیرد. هرگاه از اندازه گیری تک خوانده استفاده شده باشد امکان جایگذاری انحراف معیار متغیرهای X و Y در برنامه میسر میباشد (Figure 1). ویژگی عملکرد اندازه گیری ، ممکن است از روش دیگری تعیین شده باشد. اگر انحراف معیار در خانه های K19 و K20 وارد شده باشد مجذور نسبت آنها به عنوان λ نشان داده میشود.

Duplicate Y:s (Y/N):	Y				
Duplicate X:s (Y/N):	Y				
Student's tdep:	0,599				
α %	5		Pairs	Obs.	
Accepted samples	60	Missing Y:	0	0	
p (2-tails):	0,551	Missing X:	0	0	
No signif. difference			Coverage factor:	1,96	
Wilcoxon p:		Median or	Percentile:	97,5	
No signif. diff.	0,140	(M/S):	m	2,5 perc	97,5 perc
Y-mean:	36,64	Median:	29,8	0,73	98,26
X-mean:	36,01	Median:	31,5	0,70	96,79
Mean bias:	0,63	St dev:	8,1	SEM:	1,05
Mean rel bias %:	4,1	St dev:	18,1	SEM:	2,34
s(Y) Test :	2,2	X-min:	0,6	Y-min:	0,595
s(X) (Reference):	2,7	X-max:	98,0	Y-max:	99,0
$\lambda (s(Y)/s(X))^2$:	0,688				
s(Y):					
s(X):					

Figure 1. Calculated statistics and options for manual input.

گزارش نتایج

علاوه بر جدول شکل ۱، نتایج در سه نمودار نیز نشان داده میشوند. در نمودار پراکندگی، نتایج همگی با هم نقطه گذاری شده اند و خطوط عمودی خواننده ها را تقسیم بندی نموده اند. استفاده کننده میتواند هر دو یا یکی از خطوط رگرسیون را برای نشان داده شدن بر روی نمودار انتخاب کند. میانگین و میانه نیز نشان داده میشوند.

نمودار منحنی اختلاف به دوشکل ترسیم شده است. یکی تفاوت های مطلق و دیگری تفاوت های نسبی در ارتباط با میانگین روشها را نشان میدهد. هرگاه روش مقایسه ای، یک روش مرجع باشد، بطور مثال با مقدر کم عدم دقت، توصیه میشود مقایسه در مقابل نتایج روش مقایسه ای مرجع انجام شده و از میانگین دوروش استفاده نشود.

میانگین تورش در قسمتهای مختلف و تورش مجاز محاسبه شده $(k \times s(bias) \text{ of the total bias})$ نشان داده شده اند.

شیب منحنی $slope$ و اینترسپت $slope \text{ and intercept}$ برای رگرسیون دمیگ و رگرسیون خطی و مقادیر عدم دقت نیز نشان داده شده است.

کاربر، مجاز به انتخاب خطوط رگرسیون در نمودار میباشد.

ضریب همبستگی همواره محاسبه میگردد.

محور پراکندگی برای فواصل غلطی تنظیم شود. کاربر ممکن است تقسیم بندی نتایج را انتخاب کند. برنامه امکان انتخاب سه محدوده را برای مقادیر مقایسه ای (X) و مقادیر تست (Y) میسر ساخته است. همچنین امکان حذف انتهای توزیع روش مقایسه ای نیز وجود دارد.

تورش بین نتایج، که توسط $t\text{-test}$ و با درجه آزادی (df) مربوط محاسبه شده در هر قسمت نشان داده شده است. اگر از اندازه گیری دوتائی استفاده شده باشد، عدم دقت قسمتها نیز گزارش شده است.

درجه بندی محورهای نمودار بطور خودکار تنظیم میشوند ولی اگر نتایج اعداد بزرگی هستند، مهم است که درجه بندی از طریق افزایش مقدار حداقل محور تنظیم شود. **by increasing the minimum value of the axis.**

این تنظیم با روش استاندارد اکسل امکان پذیر میباشد. محور افقی و عمودی نمودار باید برابر باشند.

ارزیابی

وجود تفاوت معنی دار بین نتایج حاصل از دوروش اندازه گیری از طریق محاسبه t_{dep} ، در (شکل ۱) نمایش داده میشود. جدول ارزشهای وابسته به درجه آزادی (df) و میزان رد کاذب (α) ، امکان محاسبه میزان فراوانی $p\text{-value}$ را میدهد (بطور مثال هرچه $p\text{-value}$ کوچکتر باشد احتمال وجود تفاوت بین روشها بیشتر میشود). این ارزیابی بر اساس $Student's t\text{-test}$ انجام شده و باین پیش فرض است که اختلاف بین نتایج تقریباً دارای توزیع نرمال هستند. در غیر این صورت ارزیابی ممکن است دارای اعتبار نباشد.

نتایج آزمون $Wilcoxon \text{ sign rank}$ ، $studentized p\text{-value}$ و تفسیر آن نیز ارائه شده اند. اگر هر دو ارزش فراوانی، بسیار متفاوت باشند (according to Student and Wilcoxon) ارزش فراوانی مربوط به $Wilcoxon$ بیشتر قابل اعتماد است اگرچه که هر دو روش محدودیت های خود را دارند.

میانگین وانحراف معیار نتایج بیماران و میانگین تورش نسبی و تورش مطلق نیز همانطور که قبلا گفته شد محاسبه و در جدول نشان داده شده اند.

از نقطه نظر پزشکی تفاوت معنی دار آماری احتمالا مورد توجه نبوده و ارزیابی باید بیشتر بر میانگین تورش تمرکز کند. نمودارهای تفاوت، جزئیات زیادی را نشان میدهد. میانگین تورش هر قسمت **partition** و کل خواننده ها نمایش داده میشوند. انحراف معیار اختلافها نیز محاسبه شده و نمایش داده میشود.

اگر خط **TREND** در هریک از نمودارهای نسبی و یا مطلق تفاوتها افقی باشد، نشان دهنده تفاوت پیوسته در محدوده اندازه گیری است. در این حال تورش در تخمین محدوده اندازه گیری مفید و قابل استفاده است. اگر این تفاوت پیوسته نباشد در اینحال تخمین تورش در قسمتهای مختلف نتایج بر اساس روش مقایسه ای می تواند انجام شود. این نکته در جدول به عنوان **“uncertainty profile”** محدود دیده میشود.

همانطور که در شکل ۲ نشان داده شده است تا ۳ ردیف (غلظتی) را میتوان در خانه های آبی تعریف کرد. کاربر میباید محدوده گزارش را در خانه های **K22** و **M24** تعریف کند. حدود قسمتها در تمام نمودارها با خطوط عمودی نشان داده شده اند.

تعداد مشاهدات در هر قسمت ، تعدادی است که توسط روش مقایسه ای در آن محدوده تعیین شده است.

Partitioning of Reference results				Number:	Mean bias:	Bias %	s(Y)	s(X)	t _{dep}	p-value	Signif.
Low:	0,5	to	5,0	25	0,1	5,5	0,1	0,1	1,600	0,123	NS
Mid:	5	to	40,0	9	0,1	-3,4	1,6	3,5	0,084	0,935	NS
High:	40	to	98,0	26	1,4	1,9	3,3	3,5	0,568	0,575	NS

Figure 2. Partitioning of the measuring interval into three tiers

Regression:			OLR		
	Deming	$\pm U$			$\pm U$
Slope:	۱.۰۳	۰.۰۳	Slope:	۰.۹۹	۰.۰۳
Intercept:	-۰.۲۸	۱.۵۸	Intercept:	۰.۸۹	۱.۵۳
Coeff determ (r^2):	۰.۹۴۸		Display (Y/N):	N	
Display (Y/N):	Y		Displ equal line (Y/N):	Y	

Figure ۳. Characteristics of the regression functions and correlation coefficient. If a "Y" is entered the regression line for the Deming and Ordinal regression functions, respectively, will be shown in the scattergram.

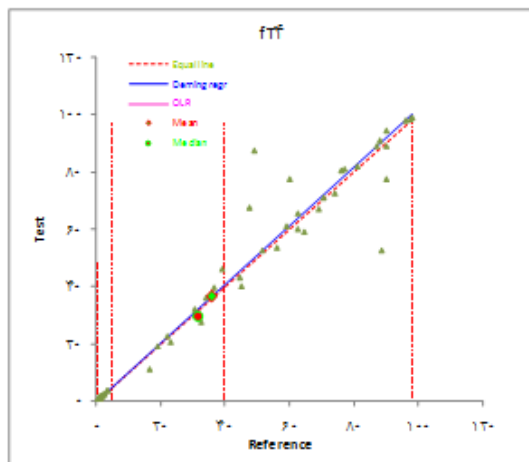


Figure ۴. Scattergram with the partitions displayed. The equal line (dotted red) and the Deming (solid blue) regression lines are shown. Note the large number of observations in the low partition and the gap between about five and ۱۷ units and the wide spread of results in the high partition.

بررسی چشمی نمودار پراکندگی، از نظر یافتن خوانده های دور افتاده احتمالی بسیار مهم است. شناسایی این خوانده توسط نمودارهای متفاوت تسهیل شده و حداکثر اختلاف بین خوانده های دوتائی را میتوان در ستون های H و I شناسایی نمود. هرگاه خوانده ای دورافتاده تشخیص داده شد به راحتی میتوان آنرا از جدول حذف نمود.

اگر تورش مطلق و یا نسبی "پیوسته constant" وجود نداشته باشد استفاده از پارامترهای رگرسیون در تخمین تورش اهمیت مییابد. اگر تورش کلینیکی معنی داری مشاهده شود روش ذکر شده برای ایجاد کالیبراسیون ممکن است کافی نبوده و در این حال برای محاسبه رگرسیون تعداد نمونه های بیشتری باید اندازه گیری شده و کالیبراسیون ۲ نقطه ای با نمونه بیماری که دارای مقادیر غلظتی معلوم و قابل ردیابی بوده و توسط روش مرجع تعیین شده مورد استفاده قرار گیرد.