

## اندازه نمونه و طول تست برای تخمین پارامتر مورد و تخمین پارامتر تست

### چکیده

مدل تئوری پاسخ مورد (IRT) شامل یک یا چند پارامتر در مدل می باشد. این پارامترها نامعلوم هستند، بنابراین لازم است تا آنها را پیش بینی نماییم. این مقاله قصد دارد تا ۱) اندازه نمونه ( $N$ ) را بر روی پایداری پارامتر مورد تعیین کند و ۲) آزمون طول ( $n$ ) بر روی پایداری پارامتر تخمینی آزمایش شونده را تعیین کند ۳) اثر مدل بر روی پایداری مورد و پارامتر مورد بررسی را تعیین نماید ۴) اثر اندازه نمونه و طول تست بر روی پایداری مورد و تخمین های پارامتر آزمایش شونده را کشف کند ۵) اثر اندازه نمونه، طول تست، و مدل بر روی پایداری مورد و تخمین های پارامتر آزمایش شونده. این مقاله یک مطالعه شبیه سازی است که در آن شبیه سازی نمونه ویژگی پنهان ( $\theta$ ) مشتق میشود از یک جمعیت نرمال استاندارد  $N(0,1)$ ، با یک اندازه نمونه مشخص ( $N$ ) و طول تست ( $n$ ) با مدل های 1PL, 2PL و 3PL با استفاده از Wingen. تحلیل آیتم با استفاده از رویکرد تست نظریه کلاسیک و نظریه آزمون مدرن انجام شد. تئوری پاسخ مورد و داده ها از طریق نرم افزار R با بسته مورد تحلیل شدند. نتایج نشان دادند که هرچه اندازه نمونه بزرگتر ( $N$ ) باشد، پارامتر تخمین زده شده پایدارتر خواهد بود. برای تست طول، که هرچه طول آزمون ( $n$ ) بیشتر باشد، پارامتر تخمین زده شده ( $\theta$ ) پایدارتر است.

کلید واژه ها: تئوری پاسخ مورد، پایداری مورد، اندازه نمونه، طول آزمون، وینگن

## مقدمه

در توسعه دانش، چندین رشته بوجود آمده اند، در زمینه آموزش و روانشناسی در مطالعه روش های اندازه گیری و راه حل آنها به یک رشته خاص بنام تئوری آزمون توسعه یافته است. راه حل های خاص و فرموله سازی روشها برای غلبه بر مسائل و اشباع آنها.

تئوری پاسخ مورد بطور قوی در اندازه گیری آموزشی در دهه ۱۹۷۰ و دهه ۱۹۹۰ استفاده شد تا یک ابزار آماری برای مدل سازی داده های پاسخ گسسته چند متغیره شود. IRT میتواند در زمینه های گوناگونی از آموزش، روان شناسی، اقتصاد، و آمار نگاری گرفته تا تحقیقات پزشکی مورد استفاده قرار گیرد. مزایای IRT این است که دارای ویژگی های روان سنجی قوی است.

تئوری پاسخ مورد یک تئوری آزمون جدید است که چگونگی ساختن استنباط ها درباره مشخصات یا توانایی های آزمون گیرندگانی که بر اساس داده های پاسخ شرکت کنندگان به موارد قابل مشاهده نیستند را توضیح میدهد. این تئوری براساس دو فرضیه می باشد یکی عملکرد امتحان دهنده بر روی مورد آزمون که قابل پیش بینی است (یا قابل توضیح است) از طریق مجموعه ای از عوامل که توانایی نامیده میشوند (صفت، صفت پنهان، یا توانایی) و دومی ارتباط بین عملکرد مورد آزمون دهندگان و مجموعه توانایی ها بر اساس عملکرد مورد و قابل توصیف شدن توسط یک تابع رو به بالا یکنواخت که تابع مشخصه مورد یا منحنی مشخصه مورد (ICC) نامیده میشود. علاوه، طبق گفته Stark سختی مورد نیز باید در نظر گرفته شود به منظور ایجاد یک آزمون با مجموع توزیع نمرات مورد نظر.

چندین مدل پاسخ مورد وجود دارد که در آنها همه مدل های IRT حاوی یک یا چند پارامتر مورد و یک یا چند پارامتر امتحان دهنده می باشند. این مقاله بر روی مدل پاسخ مورد با یک پارامتر آزمون دهنده تمرکز دارد. این پارامترها معلوم نیستند بنابراین لازم است تا حدس زده شوند. تخمین توسط شبیه سازی داده های ایجاد شده با استفاده از برنامه WinGen برای ۵ پاسخ با یک اندازه نمونه و طول آزمون معین انجام میشود. علاوه بر آن، داده های شبیه سازی با استفاده از نرم افزار R شبیه سازی میشوند جایی که مرحله تحلیل بر اساس همبستگی و معیار RMSD (خطای جذر میانگین مربعات) می باشد تا پایداری پارامترهای تخمین زده شده را ببینیم.

## روش

در یک مطالعه شبیه سازی، شبیه سازی نمونه های ویژگی پنهان مشتق می شوند از جمعیت نرمال استاندارد ~  $N(0,1)$ ، با یک اندازه نمونه ( $N$ ) و طول آزمون ( $n$ ) مشخص با مدل های 1PL, 2PL و 3PL با استفاده از Wingen. تحلیل مورد با استفاده از رویکرد تئوری آزمون کلاسیک (CTT) و آزمون تئوری پاسخ مورد مدرن (IRT) انجام شد. سپس داده ها از طریق نرم افزار R با بسته ITM تجزیه و تحلیل شدند.

## نتیجه و بحث

### ۱- اثر اندازه نمونه بر روی پایداری تخمین های پارامتر مورد

با اشاره به نظری که توسط Nunnally و همکاران پیشنهاد شد، برای مشاهده اثر اندازه نمونه بر پایداری تخمینهای پارامترهای مورد، تعداد موردها با مراجعه به قانون انگشت شست تعیین میشود (طول تست)  $n=40$  و اندازه نمونه مختلف است، یعنی  $N = 200, 400, 1000$ . هر کدام از این متغیرها ۵ بار تکرار شدند با استفاده از مدل 1PL و در همبستگی و RMSD بصورت خلاصه در جدول بصورت زیر ارائه شدند:

جدول ۱- پارامترهای همبستگی و RMSD

تکرار	همبستگی			RMSD پارامتر		
	200	400	1000	200	400	1000
1	0.991	0.9960	0.9992	0.354	0.10437	0.04828
2	0.992	0.9950	0.9992	0.422	0.11152	0.06265
3	0.991	0.9979	0.9993	0.282	0.13807	0.08213
4	0.993	0.996	0.9986	0.212	0.11625	0.0636
5	0.993	0.9961	0.999	0.314	0.13640	0.07398

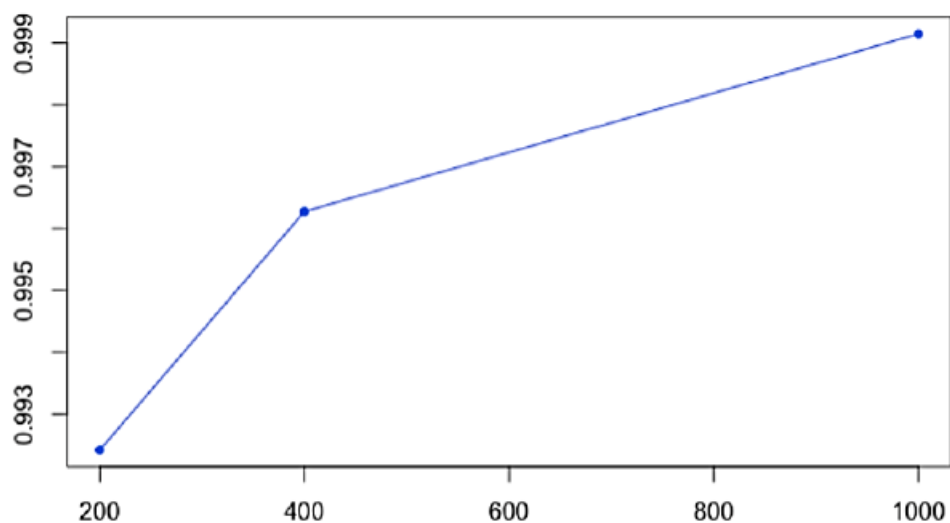
منبع: داده های پردازش شده با برنامه کاربردی Wingen

جدول ۲- پارامترهای همبستگی و RMSD

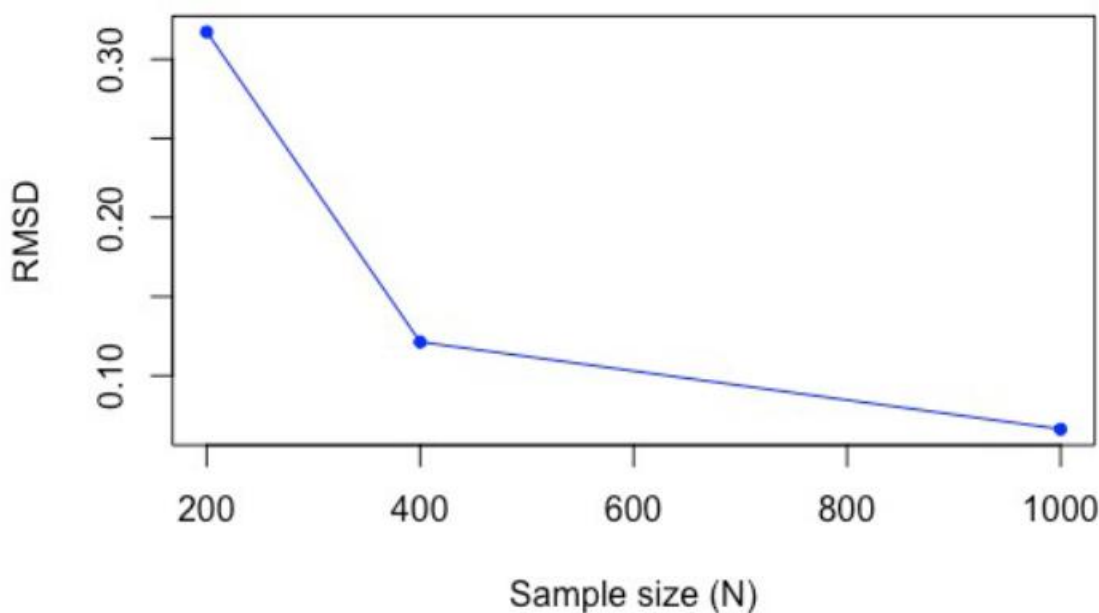
نمونه	(N)	b	RMSD
200		0.9924202	0.3172997
400		0.9962736	0.1213278
1000		0.9991467	0.06614813

منبع: داده های پردازش شده با برنامه کاربردی Wingen

براساس جدول خلاصه بالا، طرح اندازه نمونه زیر با مقادیر همبستگی بصورت زیر شکل می گیرد.



شکل ۱. گراف نتایج همبستگی از مقدار b



شکل ۲. گراف نتایج همبستگی RMSD

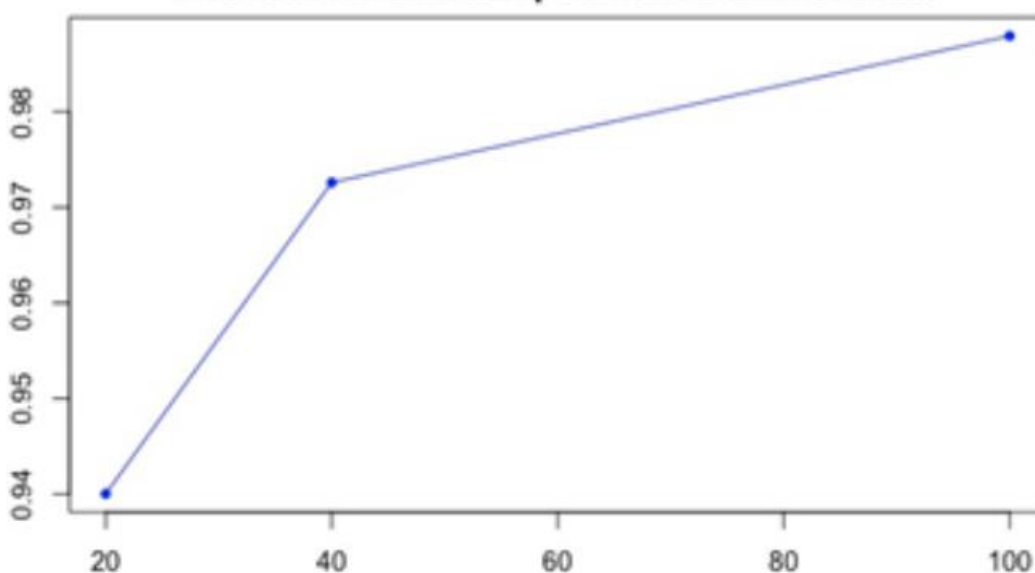
بر اساس شکل ۱ و ۲ در بالا، نشان میدهد که هرچه اندازه نمونه ( $N$ ) بزرگتر باشد، مقدار همبستگی بیشتر است، در حالی که مقدار RMSD در حال کوچکتر شدن می باشد. بنابراین اندازه نمونه ( $N$ ) بر پایداری پارامترهای تخمین زده شده اثر می گذارد، یعنی هرچه اندازه نمونه بزرگتر باشد، پارامتر تخمینی پایدارتر است.

## ۲-۲ اثر طول آزمون بر پایداری تخمین های پارامتر آزمون دهنده

به منظور تعیین اثر طول آزمون ( $n$ ) بر روی پایداری پارامتر آزمون دهنده ( $\theta$ ) در مدل منطقی دو پارامتری (2PL)، تعداد آزمون دهندگان (اندازه نمونه)  $N=1000$  تعیین شد و طول آزمون متغیر بود یعنی  $n = 20, 40, 100$ ، هر متغیر ۵ بار تکرار شد. برای مشاهده پایداری پارامتر آزمون دهنده تخمینی ( $\theta$ )، یک ضریب همبستگی و RMSD (اختلافهای مجذور میانگین ریشه) یا RMSE (خطای مجذور میانگین ریشه) استفاده شدند. نتایج همبستگی و RMSD از پارامترهای آزمون دهنده را میتوان بصورت زیر مشاهده کرد.

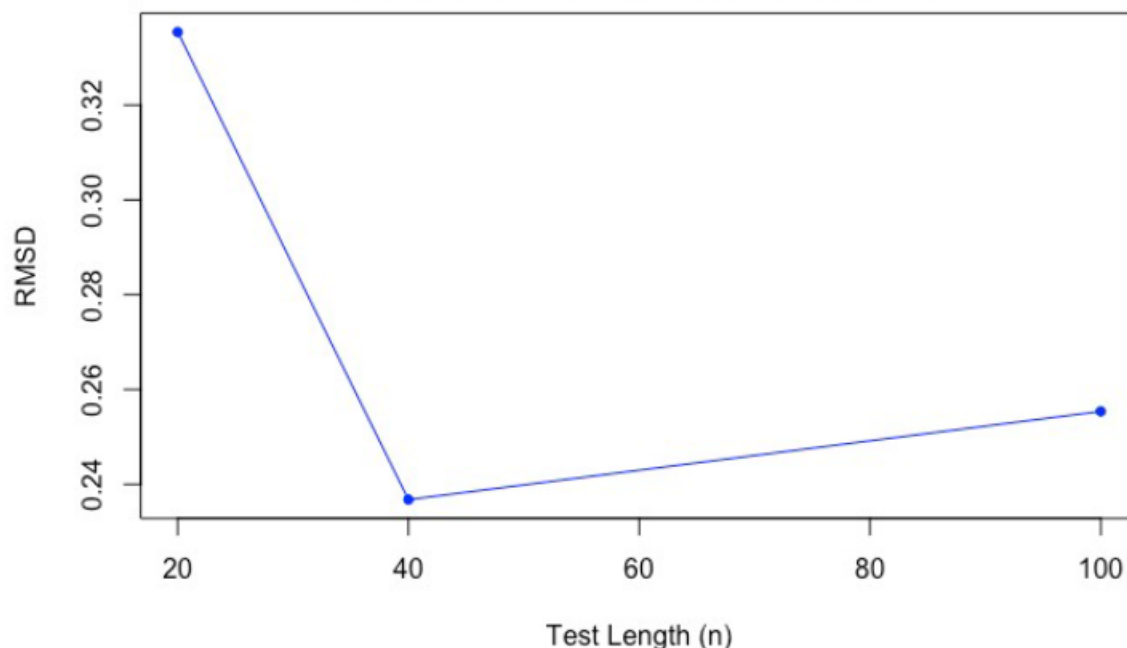
جدول ۳. نتایج همبستگی و پارامتر RMSD

طول آزمون	پارامتر آزمون دهنده	
	همبستگی	RMSD
20	0.9400224	0.3354105
40	0.9726023	0.2367724
100	0.9879318	0.2553879



شکل ۳. گراف همبستگی پارامترهای آزمون دهنده

از نمودار همبستگی پارامترهای آزمون دهنده بالا، میتوان استنتاج کرد که طول آزمون ( $n$ ) با اندازه نمونه ( $N$ ) بر مقدار واقعی  $(\theta)$  اثر خواهند داشت که هرچه طول آزمون بیشتر باشد، توانایی تخمین زده شده اندازه نمونه بالاتر خواهد بود. موضوع مشابهی نیز توسط Xing و Hambleton بیان شد که هرچه طول آزمون بیشتر باشد، قابلیت اطمینان نیز بیشتر خواهد بود.



شکل ۳. گراف RMSD پارامترهای آزمون دهنده

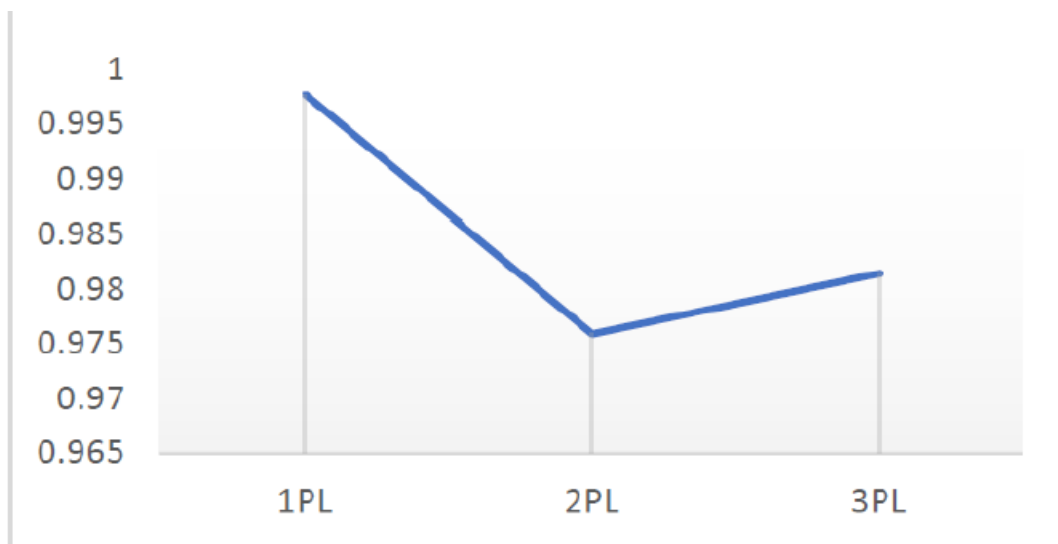
جدول ۲ با شکل ۳ نشان میدهند که طول آزمون (n) بر درستی مقدار  $(\theta)$  اثر می گذارد. مشاهده می کنیم که مقدار RMSD کوچکتر است وقتی که طول آزمون بزرگ است، یعنی طول آزمون ۴۰ است، اگرچه وقتی طول آزمون ۱۰۰ میشود، افزایش مقدار RMSD خیلی زیاد نیست. چنانکه طول آزمون (n) بر پایداری پارامتر آزمون دهنده  $(\theta)$  اثر می گذارد، یعنی هرچه طول آزمون (n) بزرگتر باشد، پارامتر تخمینی  $(\theta)$  پایدارتر خواهد بود.

۳- اثر مدل بر روی پایداری آیت‌ها و پارامترهای آزمون دهنده

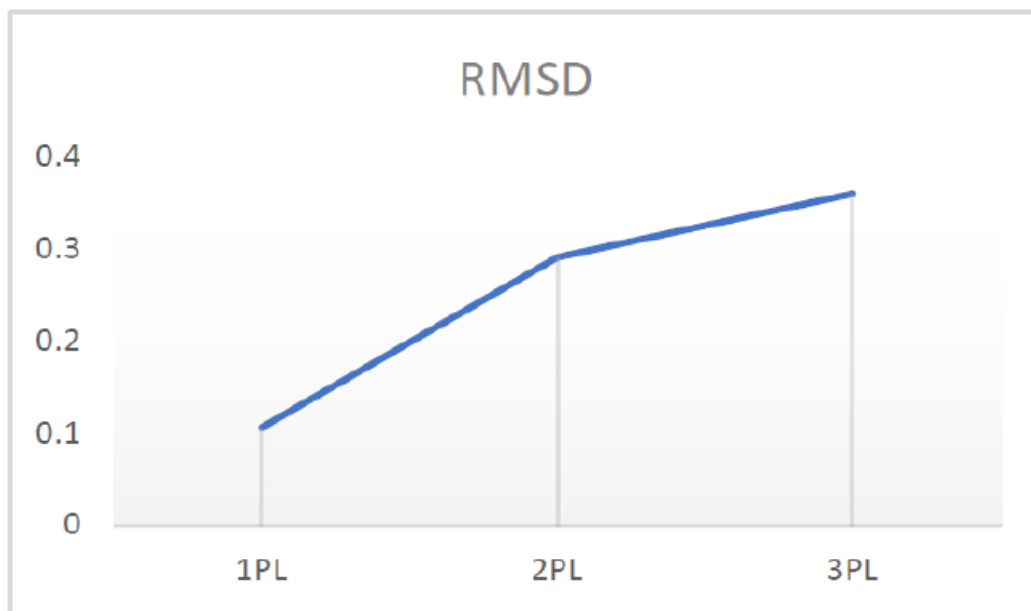
با شبیه سازی انجام شده با برنامه WinGen با ۵ تکرار با هر اندازه نمونه (N) ۴۰۰ و طول آزمون (n) ۴۰ آیت‌ها با ایجاد داده های WinGen با ۳ مدل (1PL, 2PL, 3PL) نتایج همبستگی بدست می آیند. در هر مدل پارامترهای آزمون دهنده بصورت زیر هستند:

جدول ۴. همبستگی b و RMSD با مدل

پارامتر	همبستگی	RMSD
1PL	0.9978277	0.1064805
2PL	0.9759783	0.2912042
3PL	0.9814975	0.3604854



شکل ۵. گراف نتایج همبستگی  $b$  با مدل



شکل ۶. گراف نتایج همبستگی  $b$  با مدل

بر اساس گراف نتایج RMSD با مدل، میتوان مشاهده نمود که مقدار RMSD بر خلاف مقدار پارامتر  $b$  در مدل بیشتر خواهد بود اگر مدل های بیشتری بکار گرفته شوند. یعنی 3 PL.

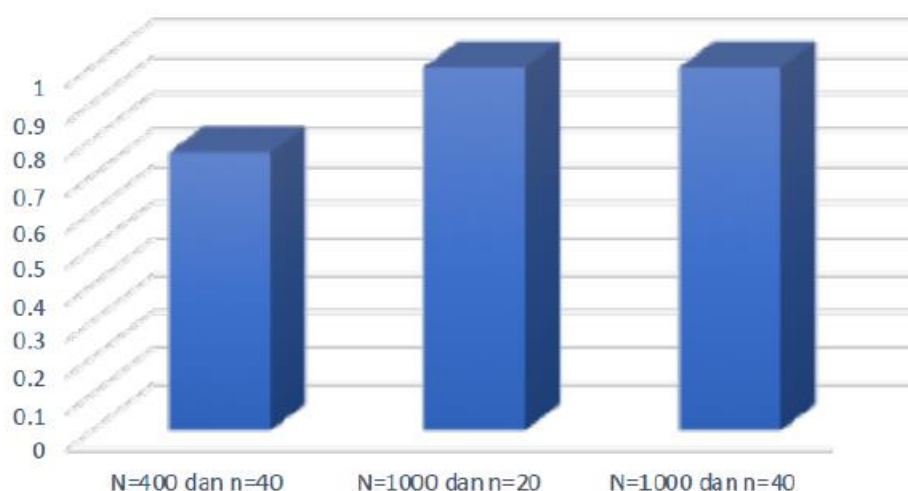
۴-تاثیر اندازه نمونه و طول آزمون بر روی پایداری مورد و تخمین های پارامتر آزمون دهنده



با انجام شبیه سازی های مختلف، یعنی با افزایش تعداد اندازه نمونه ( $N$ ) و طول آزمون ( $n$ ) با مدل PL 1 آنگاه داده ها از طریق Wingen با توزیع نرمال (0.1) تولید میشوند.

جدول ۵. نتیجه پارامتر همبستگی  $b$

N and n	b
N=400 dan n=40	0.764721
N=1000 dan n=20	0.9980098
N=1000 dan n=40	0.9982648

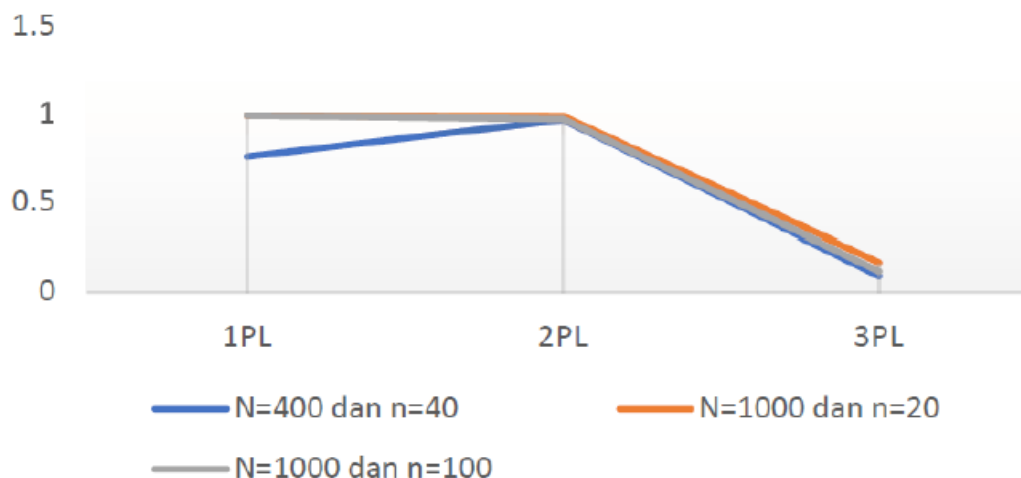


شکل ۷. گراف نتیجه همبستگی  $b$  با مدل

از گراف بالا، میتوان دید که یک اندازه نمونه بزرگ ( $N$ ) به تخمین پارامتر  $b$  اضافه خواهد شد، در حالی که طول آزمون ( $n$ ) که دارای تاثیر زیادی می باشد اما بر روی تخمین مورد پارامتر  $b$  بزرگ نیست.

۵-تاثیر اندازه نمونه، طول آزمون، و مدل بر روی پایداری مورد و تخمین های پارامتر آزمون دهنده

برای تعیین اثر تعداد  $N$  و  $n$  و انتخاب مناسب ترین مدل برای رسیدن به یک مقدار  $b$  بالا، یک شبیه سازی بر اساس این سه جنبه انجام شد. در زیر نتیجه کالبراسیون خروجی wingen پس از تکرار آن ۵ بار برای هر مدل آورده شده است.



شکل ۸. گراف نتایج تکرار پارامتر b

گراف بالا نشان میدهد که مدل 1PL اثر تعداد اندازه نمونه ( $N$ ) و طول آزمون را بر روی مقدار  $b$  نشان میدهد. میتوان مشاهده کرد که اگر  $N$  کم باشد و  $n$  زیاد باشد، مقدار همبستگی کوچک است. در حالیکه اگر  $N$  بزرگ باشد و  $n$  کوچک باشد، آن یک همبستگی بزرگ را منجر میشود. در مدل 2PL، اندازه نمونه ( $N$ ) و طول آزمون زیاد ( $n$ ) یک مقدار همبستگی بالا ایجاد میکند. اما در مدل 3PL میتوان دید که تعداد زیاد  $N$  و کم  $n$  یک مقدار بزرگ  $b$  خواهد داد.

### نتیجه گیری

پایداری پارامتر مورد تخمین زده شده تحت تاثیر اندازه نمونه می باشد و پایداری پارامتر آزمون دهنده ( $\theta$ ) تحت تاثیر اندازه طول آزمون می باشد. هرچه اندازه نمونه بزرگتر باشد، تخمین پارامتر مورد پایداری تر است، در حالیکه هرچه طول آزمون بیشتر باشد، توانایی تخمین زده شده اندازه نمونه بالاتر است. بنابراین برای حداکثر نمودن مقدار تخمین زده شده ( $\theta$ ) لازم است تا به اندازه نمونه، طول آزمون، و مدل مورد استفاده نگاه کنیم، چنانکه اگر میخواهید تا یک ابزار ارزیابی مناسبی را گردآوری کنید، شما باید به اندازه نمونه، طول آزمون، و مدل های 1PL, 2PL, 3PL مراجعه نمایید.