附录 1 定理 1~3 的证明

定理 1：设$\boldsymbol{\psi}=\{\psi_{1i},\psi_{2i}\}^{\backslash}$是对应参数$\delta=(\boldsymbol{\beta},s)$的所有估计方程的集合。根据估计方程理论，在$H_0$假设下，定义$\boldsymbol{A}=\mathrm{E}\begin{pmatrix}\frac{\partial\psi_{1i}}{\partial s} & \frac{\partial\psi_{1i}}{\partial\boldsymbol{\beta}}\\ \frac{\partial\psi_{2i}}{\partial s} & \frac{\partial\psi_{2i}}{\partial\boldsymbol{\beta}}\end{pmatrix}$，$\boldsymbol{B}=\mathrm{E}\begin{pmatrix}\psi_{1i}^2 & \psi_{1i}\psi_{2i}^{\backslash}\\ \psi_{1i}\psi_{2i} & \psi_{2i}\psi_{2i}^{\backslash}\end{pmatrix}$，则有$\sqrt{n}(\hat{s}-0)\xrightarrow{d}N(0,\sigma^2)$，其中$\sigma^2$是$\boldsymbol{A}^{-1}\boldsymbol{B}(\boldsymbol{A}^{-1})^{\backslash}$的第$(1,1)$项。令$\hat{V}_i=\left[(1+y_i)(1-\hat{\theta}_i)\hat{\theta}_i\right]$，进而新检验统计量为

$$S_{New}=\frac{\hat{s}}{\hat{\sigma}}=\frac{\sum_{i=1}^{n}\left(r_i-\hat{c}_i\right)}{\left(\sum_{i=1}^{n}\hat{c}_i\left(1-\hat{c}_i\right)-\hat{\boldsymbol{c}}(\boldsymbol{I}-\hat{\boldsymbol{c}})\boldsymbol{X}\left(\boldsymbol{X}^{\backslash}\mathrm{diag}(\hat{\boldsymbol{V}})\boldsymbol{X}\right)^{-1}\boldsymbol{X}^{\backslash}\hat{\boldsymbol{c}}(\boldsymbol{I}-\hat{\boldsymbol{c}})\right)^{1/2}}\sim N(0,1) \tag{A.1}$$

其中，$\hat{\boldsymbol{c}}=(\hat{c}_1,\hat{c}_2,...,\hat{c}_n)^{\backslash}$，$\hat{\boldsymbol{V}}=(\hat{V}_1,\hat{V}_2,\cdots\hat{V}_n)^{\backslash}$，$\boldsymbol{X}=(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n)^{\backslash}$是几何回归模型的协变量矩阵。

证明：首先，根据文中的方程组(12)，$\psi_{1i}=(r_i-c_i-s)$，$\psi_{2i}=\left(y_i\theta_i+\theta_i-1\right)\boldsymbol{x}_i^{\backslash}$，计算各矩阵，可得：

$$\boldsymbol{A}=\mathrm{E}\begin{pmatrix}\frac{\partial\psi_{1i}}{\partial s} & \frac{\partial\psi_{1i}}{\partial\boldsymbol{\beta}}\\ \frac{\partial\psi_{2i}}{\partial s} & \frac{\partial\psi_{2i}}{\partial\boldsymbol{\beta}}\end{pmatrix}=\mathrm{E}\begin{pmatrix}-1 & (1-\theta_i)\theta_i\boldsymbol{x}_i^{\backslash}\\ 0 & -\left[(1+y_i)(1-\theta_i)\theta_i\right]\boldsymbol{x}_i\,\boldsymbol{x}_i^{\backslash}\end{pmatrix}=\mathrm{E}\begin{pmatrix}-1 & (1-\theta_i)\theta_i\boldsymbol{x}_i^{\backslash}\\ 0 & -V_i\boldsymbol{x}_i\,\boldsymbol{x}_i^{\backslash}\end{pmatrix}$$

$$\boldsymbol{B}=\mathrm{E}\begin{pmatrix}\psi_{1i}^2 & \psi_{1i}\psi_{2i}^{\backslash}\\ \psi_{1i}\psi_{2i} & \psi_{2i}\psi_{2i}^{\backslash}\end{pmatrix}=\mathrm{E}\begin{pmatrix}c_i\left(1-c_i\right) & -\left(1-c_i\right)c_i\boldsymbol{x}_i^{\backslash}\\ -\left(1-c_i\right)c_i\boldsymbol{x}_i & \boldsymbol{x}_i^{\backslash}\left[(1+y_i)(1-\theta_i)\theta_i\right]\boldsymbol{x}_i\end{pmatrix}$$

$$=\mathrm{E}\begin{pmatrix}c_i\left(1-c_i\right) & -\left(1-\hat{c}_i\right)\hat{c}_i\boldsymbol{x}_i^{\backslash}\\ -\left(1-c_i\right)c_i\boldsymbol{x}_i^{\backslash} & \boldsymbol{x}_i^{\backslash}V_i\boldsymbol{x}_i\end{pmatrix}$$

其中，$V_i=\left[(1+y_i)(1-\theta_i)\theta_i\right]$。

其次，将极大似然估计$\hat{\boldsymbol{\beta}}$代入上式，得到$\hat{\boldsymbol{A}}=\begin{pmatrix}-1 & \frac{1}{n}\left[\hat{}(\boldsymbol{I}-\hat{\boldsymbol{c}})\right]^{\backslash}\boldsymbol{X}\\ 0 & -\frac{1}{n}\left(\boldsymbol{X}^{\backslash}\mathrm{diag}(\hat{\boldsymbol{V}})\boldsymbol{X}\right)\end{pmatrix}$，其中，$\left[\hat{\boldsymbol{c}}(\boldsymbol{I}-\hat{\boldsymbol{c}})\right]^{\backslash}$是两个向量的逐点乘积。而其矩阵的逆$\boldsymbol{A}^{-1}$的估计为$\hat{\boldsymbol{A}}^{-1}$，即

$$\hat{\boldsymbol{A}}^{-1}=\begin{pmatrix}-1 & -\left[(\boldsymbol{I}-\hat{\boldsymbol{c}})\hat{\boldsymbol{c}}\right]^{\backslash}\boldsymbol{X}\left(\boldsymbol{X}^{\backslash}\mathrm{diag}(\hat{\boldsymbol{V}})\boldsymbol{X}\right)^{-1}\\ 0 & -n\left(\boldsymbol{X}^{\backslash}\mathrm{diag}(\hat{\boldsymbol{V}})\boldsymbol{X}\right)^{-1}\end{pmatrix}$$

类似地有，

$$\boldsymbol{B}=\frac{1}{n}\begin{pmatrix}\sum_{i=1}^{n}\hat{c}_i\left(1-\hat{c}_i\right) & -\left[(\boldsymbol{I}-\hat{\boldsymbol{c}})\hat{\boldsymbol{c}}\right]^{\backslash}\boldsymbol{X}\\ -\boldsymbol{X}^{\backslash}\left[(\boldsymbol{I}-\hat{\boldsymbol{c}})\hat{\boldsymbol{c}}\right] & \boldsymbol{X}^{\backslash}\mathrm{diag}(\hat{\boldsymbol{V}})\boldsymbol{X}\end{pmatrix}。$$

最后，令 $\boldsymbol{D}=\left(\boldsymbol{X}^{\backslash}\operatorname{diag}(\hat{\boldsymbol{V}})\boldsymbol{X}\right)$，则可以得到：

$$
\begin{aligned}
&\hat{\boldsymbol{A}}^{-1}\hat{\boldsymbol{B}}(\hat{\boldsymbol{A}}^{-1})^{\backslash}\\
&=\frac{1}{n}\begin{pmatrix}-1 & -\left[(\boldsymbol{I}-\hat{\boldsymbol{c}})\hat{\boldsymbol{c}}\right]^{\backslash}\boldsymbol{X}\boldsymbol{D}^{-1}\\ 0 & -n\boldsymbol{D}^{-1}\end{pmatrix}\cdot\begin{pmatrix}\sum_{i=1}^{n}\hat{c}_i(1-\hat{c}_i) & -\left[(\boldsymbol{I}-\hat{\boldsymbol{c}})\hat{\boldsymbol{c}}\right]^{\backslash}\boldsymbol{X}\\ -\boldsymbol{X}^{\backslash}\left[(\boldsymbol{I}-\hat{\boldsymbol{c}})\hat{\boldsymbol{c}}\right] & \boldsymbol{D}\end{pmatrix}\cdot\begin{pmatrix}-1 & 0\\ -\boldsymbol{D}^{-1}\boldsymbol{X}^{\backslash}\left[(\boldsymbol{I}-\hat{\boldsymbol{c}})\hat{\boldsymbol{c}}\right] & -n\boldsymbol{D}^{-1}\end{pmatrix}\\
&=\frac{1}{n}\begin{pmatrix}-\sum_{i=1}^{n}\hat{c}_i(1-\hat{c}_i)+\left[(\boldsymbol{I}-\hat{\boldsymbol{c}})\hat{\boldsymbol{c}}\right]^{\backslash}\boldsymbol{X}\boldsymbol{D}^{-1}\boldsymbol{X}^{\backslash}\left[(\boldsymbol{I}-\hat{\boldsymbol{c}})\hat{\boldsymbol{c}}\right] & 0\\ n\boldsymbol{D}^{-1}\boldsymbol{X}^{\backslash}\left[(\boldsymbol{I}-\hat{\boldsymbol{c}})\hat{\boldsymbol{c}}\right] & -n\boldsymbol{D}^{-1}\boldsymbol{D}\end{pmatrix}\cdot\begin{pmatrix}-1 & 0\\ -\boldsymbol{D}^{-1}\boldsymbol{X}^{\backslash}\left[(\boldsymbol{I}-\hat{\boldsymbol{c}})\hat{\boldsymbol{c}}\right] & -n\boldsymbol{D}^{-1}\end{pmatrix}\\
&=\frac{1}{n}\begin{pmatrix}\sum_{i=1}^{n}\hat{c}_i(1-\hat{c}_i)-\left[(\boldsymbol{I}-\hat{\boldsymbol{c}})\hat{\boldsymbol{c}}\right]^{\backslash}\boldsymbol{X}\boldsymbol{D}^{-1}\boldsymbol{X}^{\backslash}\left[(\boldsymbol{I}-\hat{\boldsymbol{c}})\hat{\boldsymbol{c}}\right] & 0\\ 0 & n^2\boldsymbol{D}^{-1}\end{pmatrix},
\end{aligned}
$$

其中，第(1,1)项 $\left[\frac{1}{n}\sum_{i=1}^{n}\hat{c}_i(1-\hat{c}_i)-\left[(\boldsymbol{I}-\hat{\boldsymbol{c}})\hat{\boldsymbol{c}}\right]^{\backslash}\boldsymbol{X}\boldsymbol{D}^{-1}\boldsymbol{X}^{\backslash}\left[(\boldsymbol{I}-\hat{\boldsymbol{c}})\hat{\boldsymbol{c}}\right]\right]^{1/2}$ 是 $\boldsymbol{s}$ 的标准误，由此证明式(A.1)。

定理 2：在零假设下，基于定理 1 构造的新检验统计量渐近服从标准正态分布，即 $S_{New}\to N(0,1)$

证明：利用参数估计值 $\hat{s},\hat{\boldsymbol{\beta}}$ 代替 $s,\boldsymbol{\beta}$ 得到方差估计值，由此得到新的检验统计量形式如下：

$$
S_{New}=\frac{\sum_{i=1}^{n}\left(r_i-\hat{c}_i\right)}{\left(\sum_{i=1}^{n}\hat{c}_i(1-\hat{c}_i)-\hat{\boldsymbol{c}}(\boldsymbol{I}-\hat{\boldsymbol{c}})\boldsymbol{X}\left(\boldsymbol{X}^{\backslash}\operatorname{diag}(\hat{\boldsymbol{V}})\boldsymbol{X}\right)^{-1}\boldsymbol{X}^{\backslash}\hat{\boldsymbol{c}}(\boldsymbol{I}-\hat{\boldsymbol{c}})\right)^{1/2}}
$$

由定理 1 和 Slutskey 定理可得定理 2 成立。

定理 3：当几何回归模型(1)中不涉及协变量，且均值为常数时，$\hat{c}_i$ 是常数，新检验统计量 S_{New} 等价于得分检验统计量 S_{Score}。

证明：当没有协变量时，$\hat{c}_i$ 是一个常数，当 $\hat{\boldsymbol{c}}=\hat{c}_i,i=1,2,...,n$ 时，新检验统计量可以写成

$$
S_{New}=\frac{\sum_{i=1}^{n}\left(r_i-\hat{\boldsymbol{c}}\right)}{\left(\hat{\boldsymbol{c}}(\boldsymbol{I}-\hat{\boldsymbol{c}})-\hat{\boldsymbol{c}}(\boldsymbol{I}-\hat{\boldsymbol{c}})\boldsymbol{X}\left(\boldsymbol{X}^{\backslash}\operatorname{diag}(\hat{\boldsymbol{V}})\boldsymbol{X}\right)^{-1}\boldsymbol{X}^{\backslash}\hat{\boldsymbol{c}}(\boldsymbol{I}-\hat{\boldsymbol{c}})\right)^{1/2}}\tag{A.2}
$$

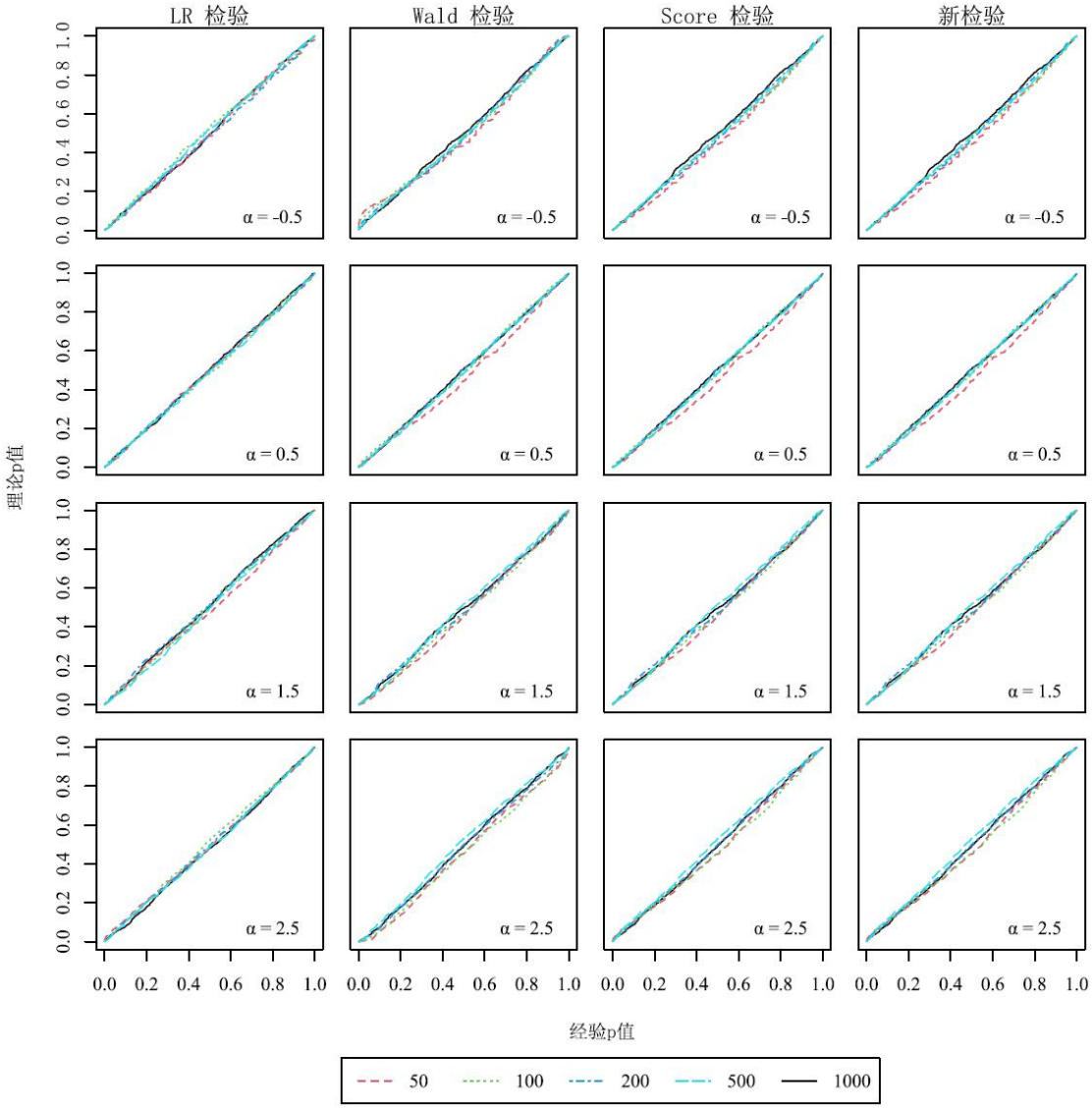
对于检验 $\phi=0$ 的 Score 统计量，当 $\hat{\boldsymbol{c}}=\hat{c}_i,i=1,2,...,n$ 时，统计量可变形为

$$
\begin{aligned}
S_{score}&=\frac{\sum_{i=1}^{n}\left(\frac{r_i-\hat{\boldsymbol{c}}}{\hat{\boldsymbol{c}}}\right)}{\left(\frac{\boldsymbol{I}-\hat{\boldsymbol{c}}}{\hat{\boldsymbol{c}}}-\frac{n}{\hat{\boldsymbol{c}}}(\boldsymbol{I}-\hat{\boldsymbol{c}})\boldsymbol{X}\left(\boldsymbol{X}^{\backslash}\operatorname{diag}(\hat{\boldsymbol{V}})\boldsymbol{X}\right)^{-1}\boldsymbol{X}^{\backslash}\hat{\boldsymbol{c}}(\boldsymbol{I}-\hat{\boldsymbol{c}})\right)^{1/2}}\\
&=\frac{\sum_{i=1}^{n}\left(r_i-\hat{\boldsymbol{c}}\right)}{\left(\hat{\boldsymbol{c}}(\boldsymbol{I}-\hat{\boldsymbol{c}})-\hat{\boldsymbol{c}}(\boldsymbol{I}-\hat{\boldsymbol{c}})\boldsymbol{X}\left(\boldsymbol{X}^{\backslash}\operatorname{diag}(\hat{\boldsymbol{V}})\boldsymbol{X}\right)^{-1}\boldsymbol{X}^{\backslash}\hat{\boldsymbol{c}}(\boldsymbol{I}-\hat{\boldsymbol{c}})\right)^{1/2}}
\end{aligned}
$$

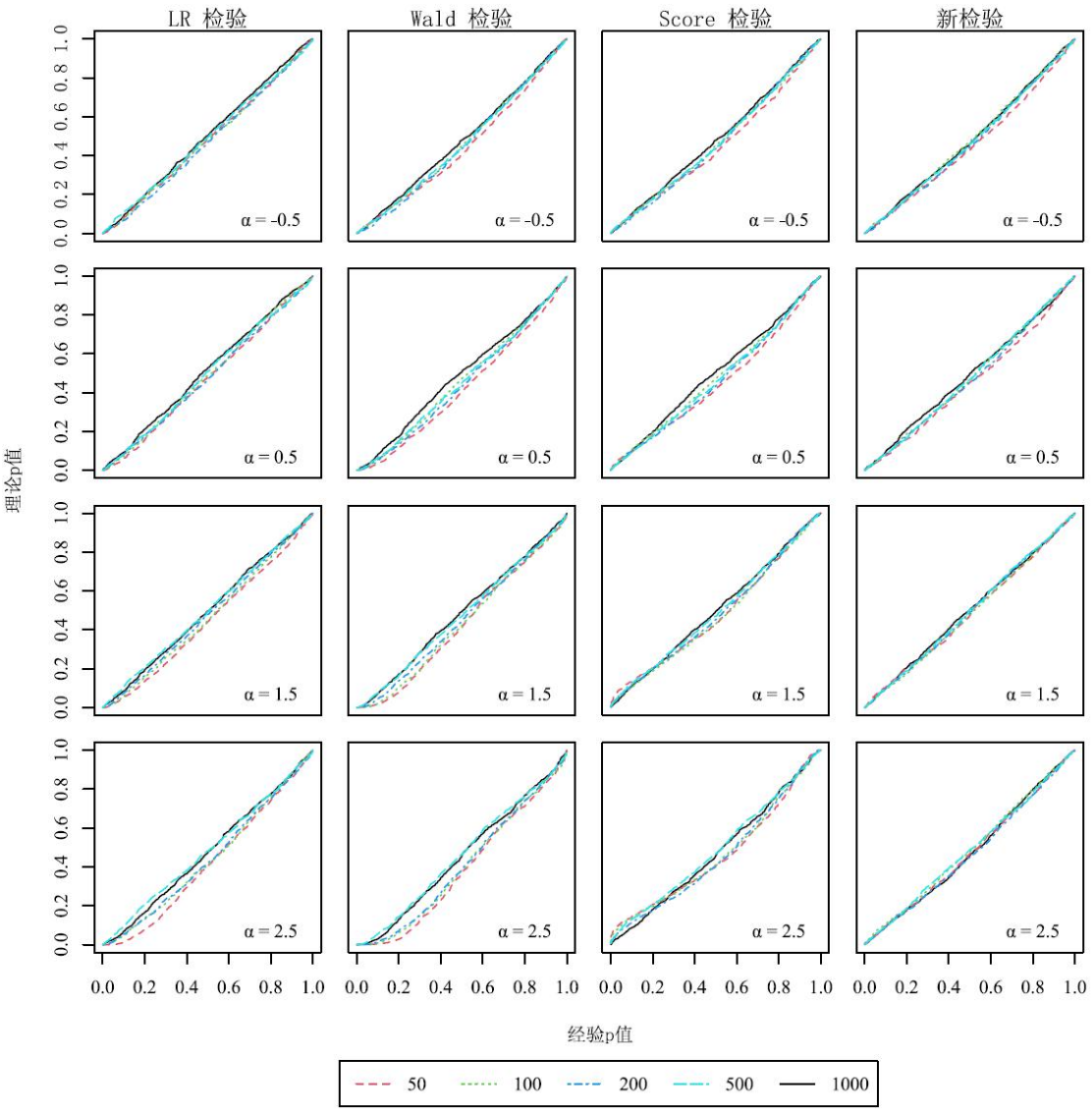
该证明过程与文献(Zou 等, 2021)类似，上式与式(A.2)完全相同，即定理证毕。

附录 2 模拟实验附图

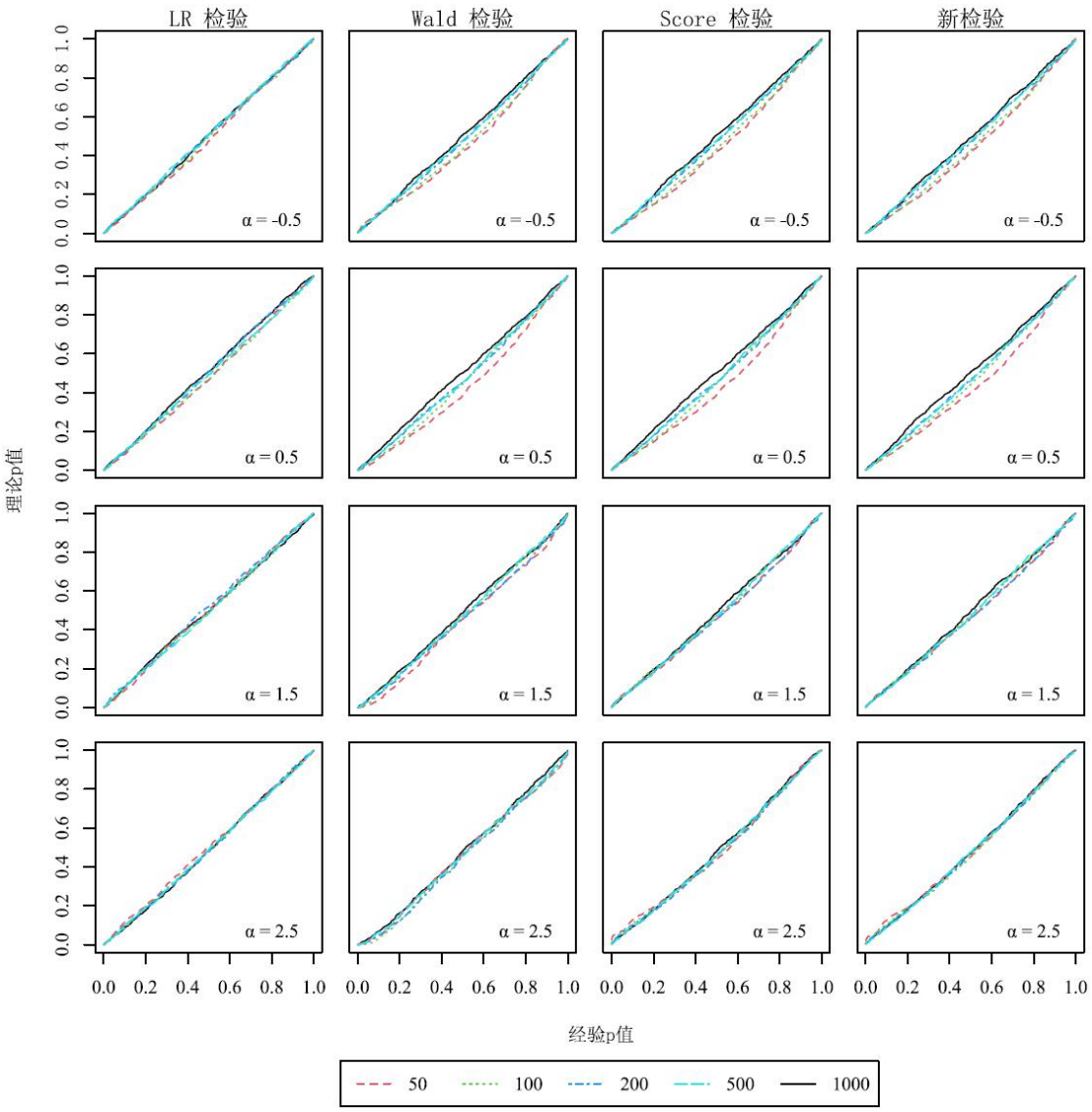
在第三章模拟实验中，零假设为真的情况下，即数据服从几何分布时，评价不同检验方法错误地拒绝零假设的概率，即考虑不同情形下不同方法对第一类错误控制水平，本文在几何回归模型下做出 3 种场景（均值是不依赖协变量的常数，均值服从正态分布的协变量，均值服从均匀分布的协变量）的 QQ 图如附图 1~3 所示。此外，当数据服从零膨胀几何分布时，考虑不同检验方法有效识别零膨胀特征的功效，即正确拒绝零假设的概率。在零膨胀几何回归模型下同样也做出 3 种场景的功效变化图如附图 4~8 所示。



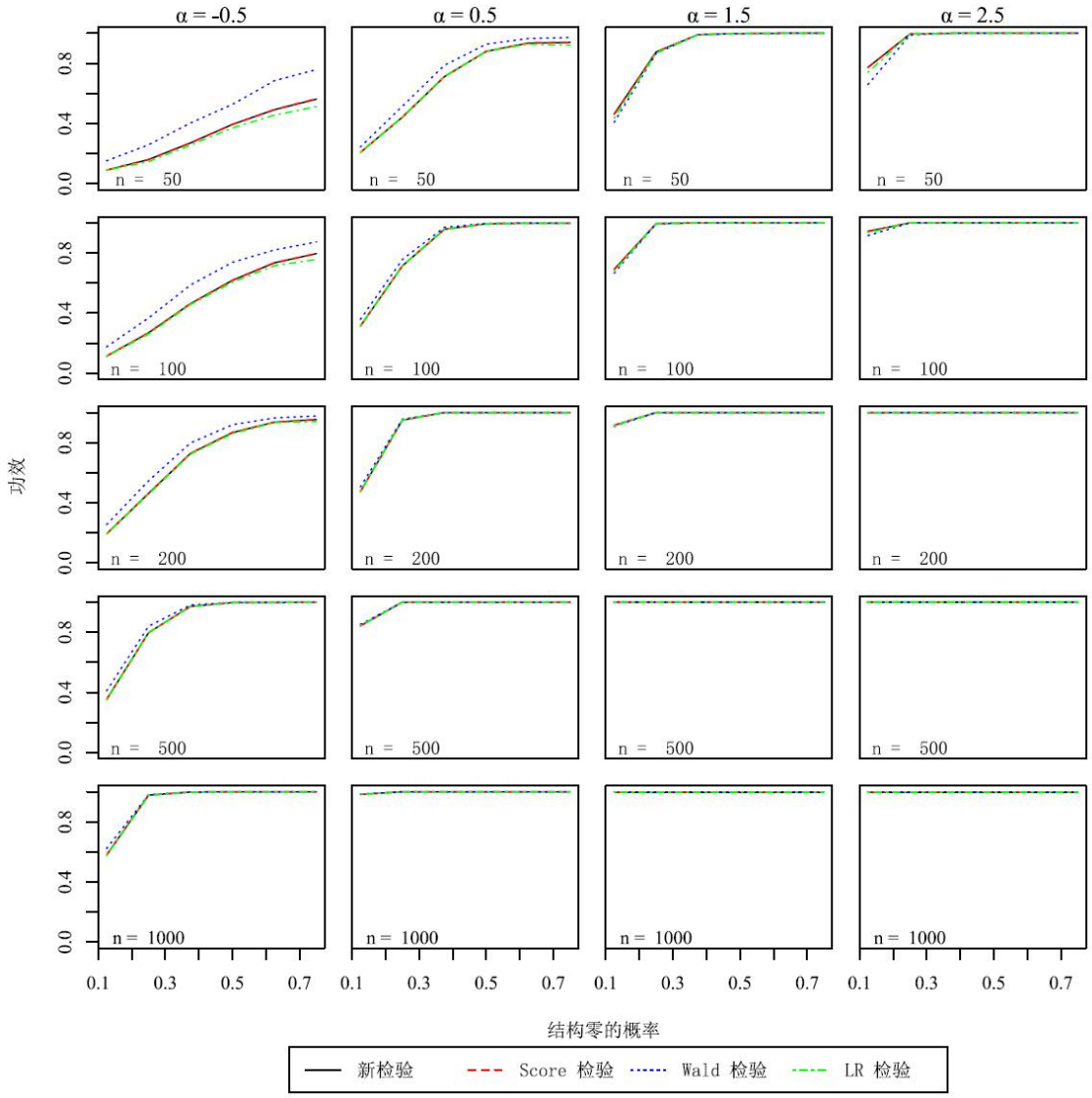
附图 1 几何回归模型下不同样本量时无协变量的 QQ 图



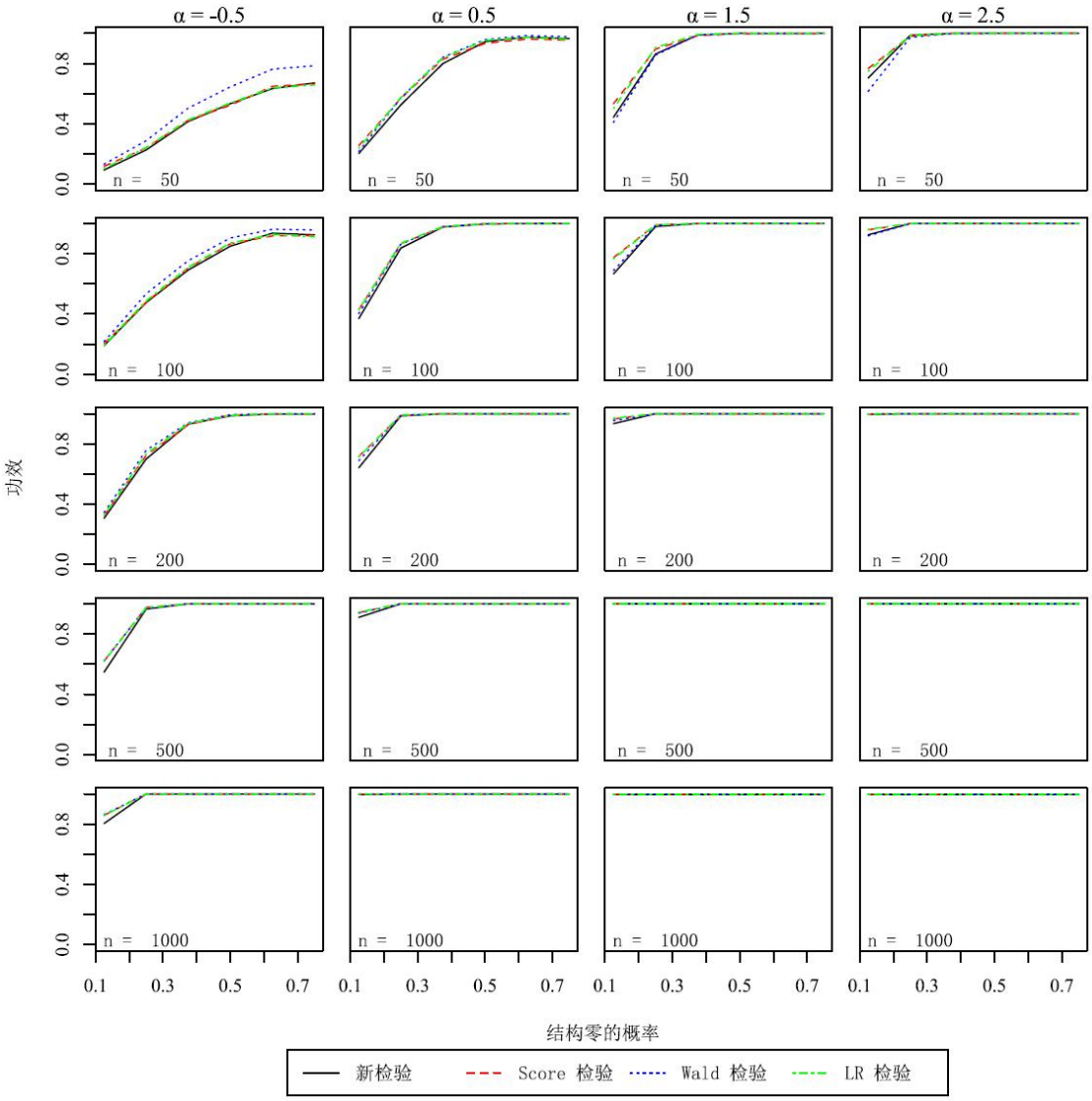
附图 2 几何回归模型下不同样本量时协变量服从正态分布的 QQ 图



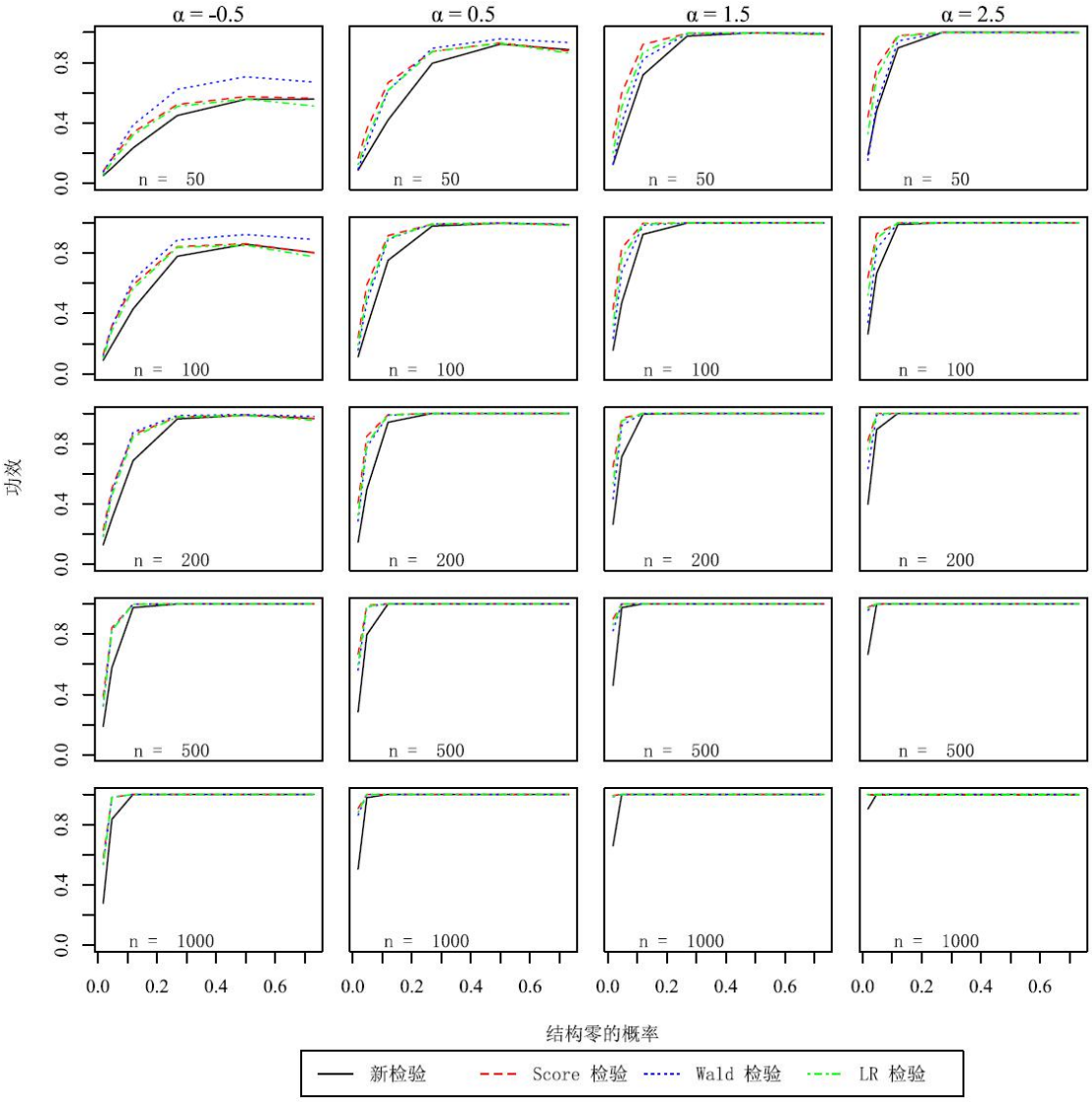
附图 3 几何回归模型下不同样本量时协变量服从均匀分布的 QQ 图



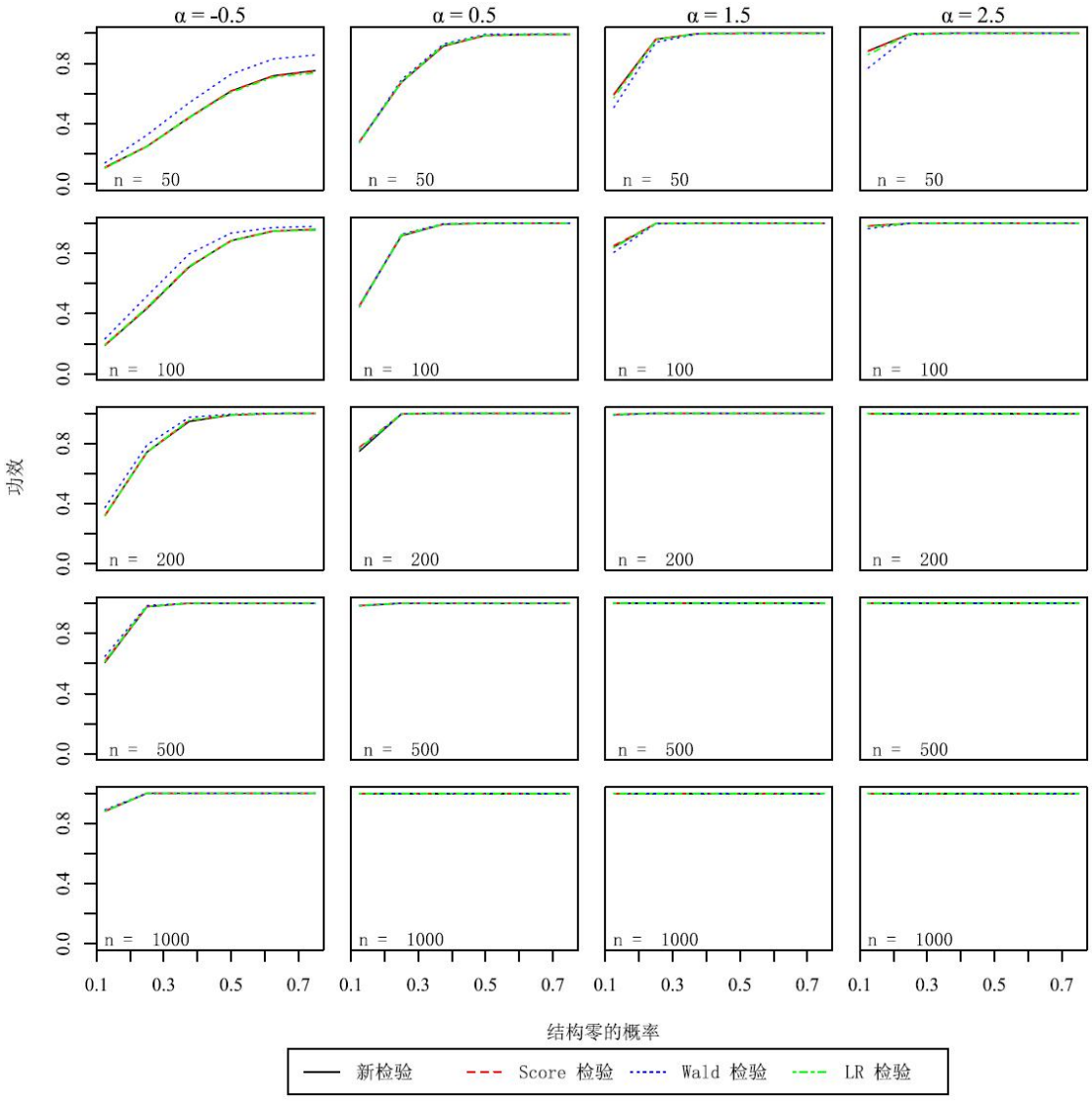
附图 4 无协变量时 ZIGR 模型零膨胀检验功效



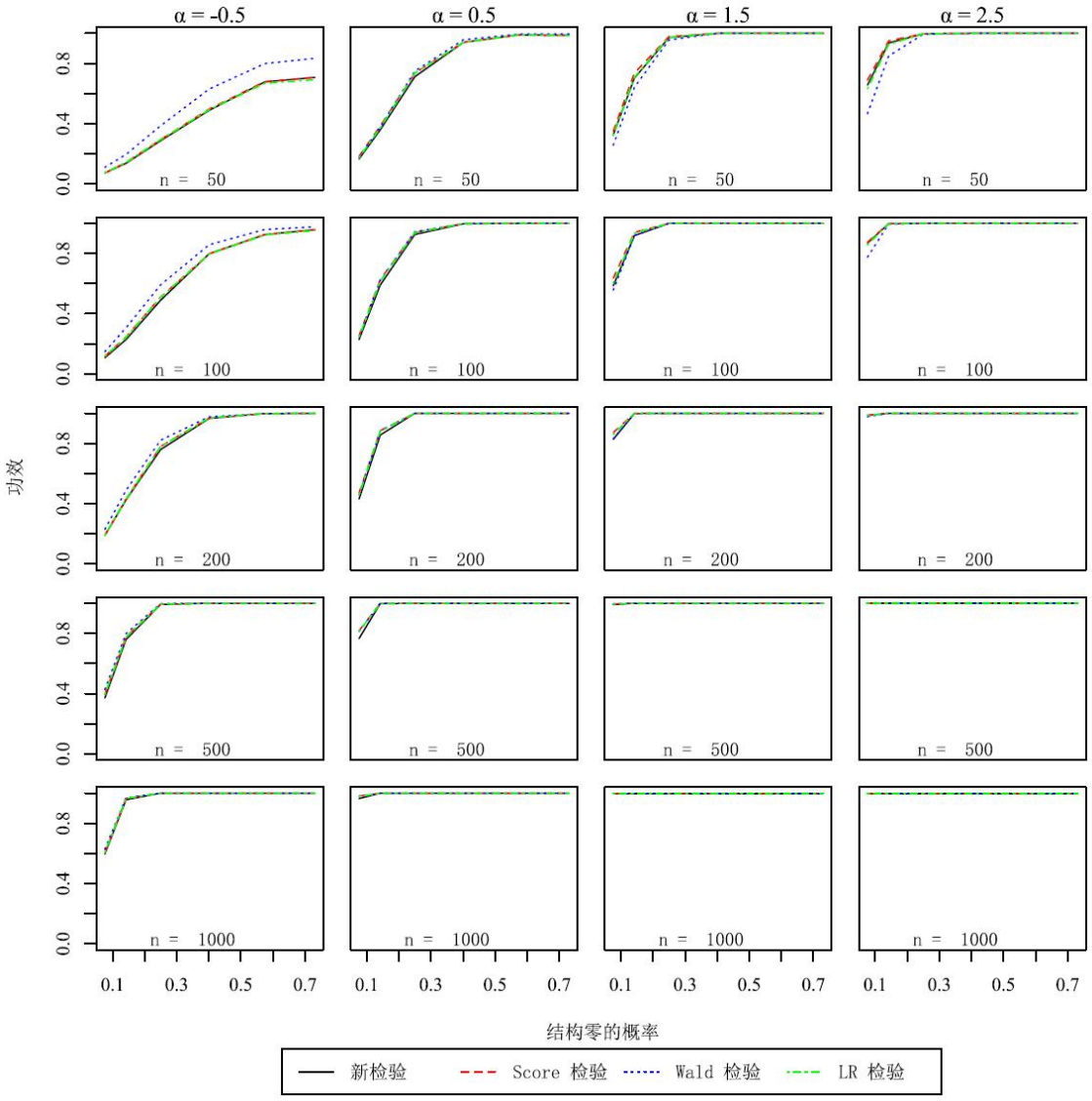
附图 5 均值的协变量服从正态分布，零膨胀参数 ϕ 无协变量时 ZIGR 模型零膨胀检验功效



附图 6 均值的协变量服从正态分布，零膨胀参数 ϕ 依赖协变量时 ZIGR 模型零膨胀检验功效



附图 7 均值的协变量服从均匀分布，零膨胀参数 ϕ 无协变量时 ZIGR 模型零膨胀检验功效



附图 8 均值的协变量服从均匀分布，零膨胀参数 ϕ 依赖协变量时 ZIGR 模型零膨胀检验功效

参考文献

[1] Zou Y, Peng Z, Cornell J, et al. A new statistical test for latent class in censored data due to detection limit[J]. Statistics in Medicine, 2021, 40(3): 779-798.