Пример.

Графическая проверка гипотезы отрицательной биномиальности против альтернативы хофмановости.

Пусть $y = (y_1, \ldots, y_n)^T$, $y_i \in \mathsf{N}_+$, y_i – н.о.р. случайные величины, $i = \overline{1,n}$.

При гипотезе Γ_1

$$y_i \overset{d}{=} NBIN(\nu_0, \tau_0), \tag{1}$$

с параметрами $\nu_0 > 0$, $0 < \tau_0 < 1$ и производящей функцией

$$\mathbf{E}\{z^{y_1}; \Gamma_1\} = \pi_1(z; \nu_0, \tau_o) = \frac{1}{\left(1 + \left(\frac{1-\tau_0}{\tau_0}\right)(1-z)\right)^{\nu_0}}. \tag{2}$$

При гипотезе Γ_2

$$y_1 \overset{d}{=} HOFM(\alpha_0, \beta_0, \gamma_0) \tag{3}$$

с параметрами $\alpha_0 > 0$, $\beta_0 > 0$, $\gamma_0 \geq 0$, $\gamma_0 \neq 1$ и производящей функцией

$$\mathbf{E}\{z^{y_1}; \Gamma_2\} = \pi_2(z; \alpha_0, \beta_o, \gamma_0) = \exp\left\{-\frac{\alpha_0}{\beta_0(1-\gamma_0)}\left((1 + \beta_0(1-z))^{1-\gamma_0} - 1\right)\right\}. \tag{4}$$

При построении "$\pi - \pi$" вероятностного графика полагаем

$$\begin{cases} R = \pi_2(z; \hat\alpha_n, \hat\beta_n, \hat\gamma_n), \\ \\ r = \pi_1(z; \hat\nu_n, \hat\tau_n), \end{cases} \tag{5}$$

где $\hat\alpha_n$, $\hat\beta_n$, $\hat\gamma_n$ состоятельные оценки параметров α_0, β_0 и γ_0 в предположении справедливости гипотезы Γ_2, а $\hat\nu_n$, $\hat\tau_n$ – состоятельные оценки параметров ν_0 и τ_0 в предположении справедливости гипотезы Γ_1.

Обратим внимание, что в (5) представление R выбрано правильно. Дело в том, что соответствующую формулу можно считать правильной для производящей функции и при объединенной гипотезе $\Gamma_1 \cup \Gamma_2$, т.к.

$$\lim_{\gamma_0 \to 1} \pi_2(z; \alpha_0, \beta_0, \gamma_0) = \pi_1(z; \nu_0, \tau_0), \tag{6}$$

где $\nu_0 = \frac{\alpha_0}{\beta_0}$, $\tau_0 = \frac{1}{1+\beta_0}$ в параметризации гипотезы Γ_2.

Таким образом, при $n \to \infty$ и справедливости гипотезы Γ_1 предельная форма кривой (5) является биссектрисой (т.е. прямой) первого квадранта.

Перейдем к определению предельной формы кривой (5) в случае справедливости гипотезы Γ_2. В качестве упомянутых выше состоятельных оценок можно, в принципе, взять оценки любого типа. Однако, будем считать, что все оценки и при Γ_1, и при Γ_2, получены по методу моментов. Выбор метода оценивания связан с тем, что для рассматриваемых статистических моделей значение оценок и их предельное значение легко получить в явном виде. Если

$$\hat f_k = \frac{1}{n}\sum_{i=1}^{n} y_i(y_i - 1)\cdots(y_i - k + 1)$$

эмпирические факториальные моменты, то при гипотезе Γ_1 оценки $\hat\nu_n$ и $\hat\tau_n$ удовлетворяют системе уравнений

$$\begin{cases} \hat{f}_1 = \hat{\nu}_n \frac{1-\hat{\tau}_n}{\hat{\tau}_n}, \\ \\ \hat{f}_2 = \hat{\nu}_n(\hat{\nu}_n+1)\left(\frac{1-\hat{\tau}_n}{1-\hat{\tau}_n}\right)^2, \end{cases} \tag{7}$$

а $\hat{\alpha}_n$, $\hat{\beta}_n$ и $\hat{\gamma}_n$ при Γ_2 удовлетворяют системе уравнений

$$\begin{cases} \hat{f}_1 = \hat{\alpha}_n, \\ \\ \hat{f}_2 = \hat{\alpha}_n^2 + \hat{\alpha}_n\hat{\beta}_n\hat{\gamma}_n, \\ \\ \hat{f}_3 = \hat{\alpha}_n^3 + 3\hat{\alpha}_n^2\hat{\beta}_n\hat{\gamma}_n + \hat{\alpha}_n\hat{\beta}_n^2\hat{\gamma}_n(\hat{\gamma}_n+1). \end{cases} \tag{8}$$

Из системы (7) получаем, что

$$\begin{cases} \hat{\nu}_n = \frac{\hat{f}_1^2}{\hat{f}_2-\hat{f}_1^2}, \\ \\ \hat{\tau}_n = \frac{\hat{f}_2}{\hat{f}_2-\hat{f}_1^2+f_1}, \\ \\ \frac{1-\hat{\tau}_n}{\hat{\tau}_n} = \frac{f_2-\hat{f}_1^2}{\hat{f}_1}. \end{cases} \tag{9}$$

При справедливости гипотезы Γ_1 и $n \to \infty$ оценки

$$\begin{cases} \hat{\nu}_n = \nu_0 + o_p(1), \\ \\ \hat{\tau}_n = \tau_0 + o_p(1), \\ \\ \frac{1-\hat{\tau}_n}{\hat{\tau}_n} = \frac{1-\tau_0}{\tau_0} + o_p(1), \end{cases} \tag{10}$$

а при справедливости гипотезы Γ_2 и $n \to \infty$

$$\begin{cases} \hat{\nu}_n = \frac{\alpha_0}{\beta_0\gamma_0} + o_p(1), \\ \\ \hat{\tau}_n = \frac{1}{1+\beta_0\gamma_0} + o_p(1), \\ \\ \frac{1-\hat{\tau}_n}{\hat{\tau}_n} = \beta_0\gamma_0 + o_p(1). \end{cases} \tag{11}$$

Таким образом, чтобы найти в плоскости $r{-}0{-}R$ каталог кривых $R = R(r; \alpha_0, \beta_0, \gamma_0)$, определяемых при Γ_2 параметрическим представлением (5), выразим $(z-1)$ через r, α_0, β_0 и γ_0, используя второе из равенств системы (5):

$$r = \frac{1}{(1+\beta_0\gamma_0(1-z))^{\frac{\alpha_0}{\beta_0\gamma_0}} + o_p(1)}. \tag{12}$$

Из (12) имеем

$$(1-z) = \frac{1}{\beta_0\gamma_0}\left(r^{-\frac{\beta_0\gamma_0}{\alpha_0}} - 1\right). \tag{13}$$

С учетом явного вида (4) производящей функции $\pi_2(z; \alpha_0, \beta_0, \gamma_0)$, подставим в правую часть равенства (13) вместо $(z-1)$ правой части первого из системы равенств (5) и получим

$$R = \exp\left\{-\frac{\alpha_0}{\beta_0(1-\gamma_0)}\left(\left\{1+\frac{1}{\gamma_0}\left(r^{-\frac{\beta_0\gamma_0}{\alpha_0}}-1\right)\right\}^{1-\gamma_0}-1\right)\right\}+o_p(1). \tag{14}$$

Таким образом, если кривая (5), реально построенная на основе данных y с использованием оценок вида (9), уклонается от биссектрисы в сторону одной из кривых вида (14), то гипотезу Γ_1 имеет смысл отклонить в пользу гипотезы Γ_2.

Надо отметить однако, что на глаз определить тип кривой вида (14) не просто. Имеет смысл ее распрямить, т.е. положить.

$$\tilde{R} = \left\{\gamma_0\left(\left\{-\frac{\beta_0(1-\gamma_0)}{\alpha_0}\ln R+1\right\}^{\frac{1}{1-\gamma_0}}-1\right)+1\right\}^{-\frac{\alpha_0}{\beta_0\gamma_0}}, \tag{15}$$

то в плоскости $r-0-R$ кривые (14) преобразуются в кривые вида

$$\tilde{R} = r + o_d(1). \tag{16}$$

Для этого в (15) вместо $\frac{\alpha_0}{\beta_0}$ и γ_0 необходимо взять их моментные оценки. Попутно заметим, что кривые (14) и преобразование (15) зависит не от индивидуальных значений α_0 и β_0, а лишь от их отношений $\frac{\alpha_0}{\beta_0}$. ■