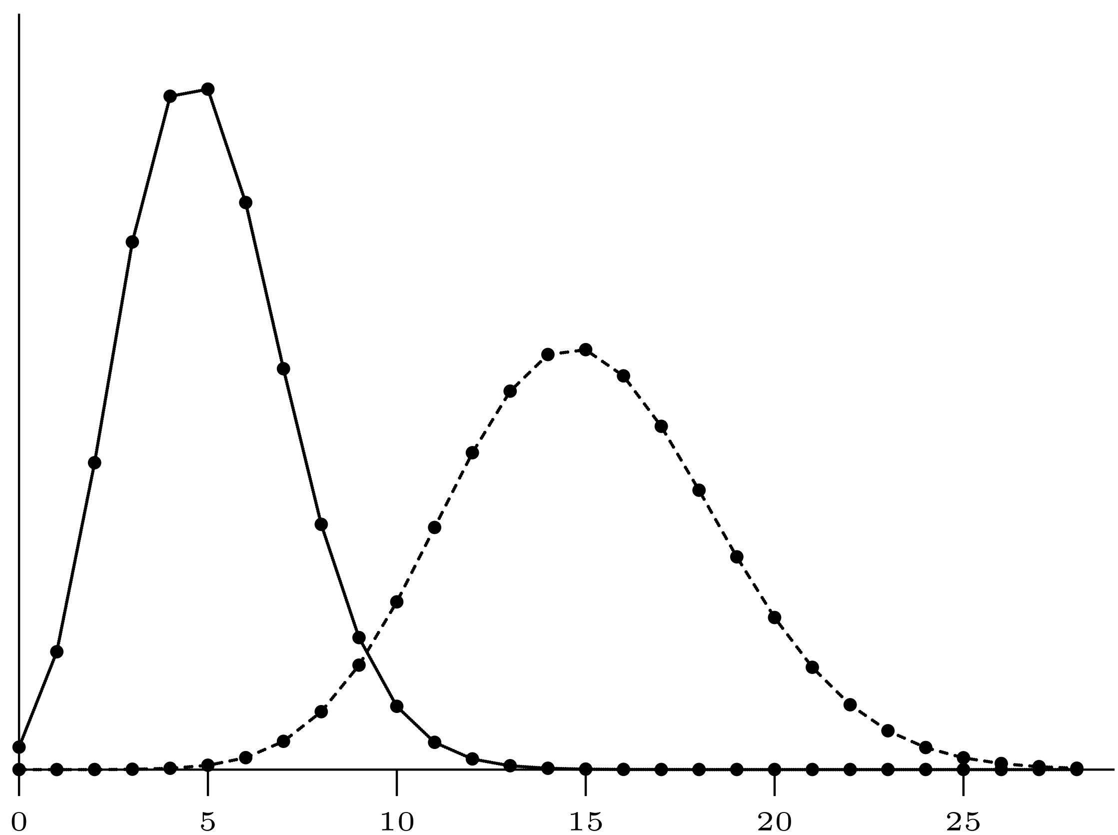
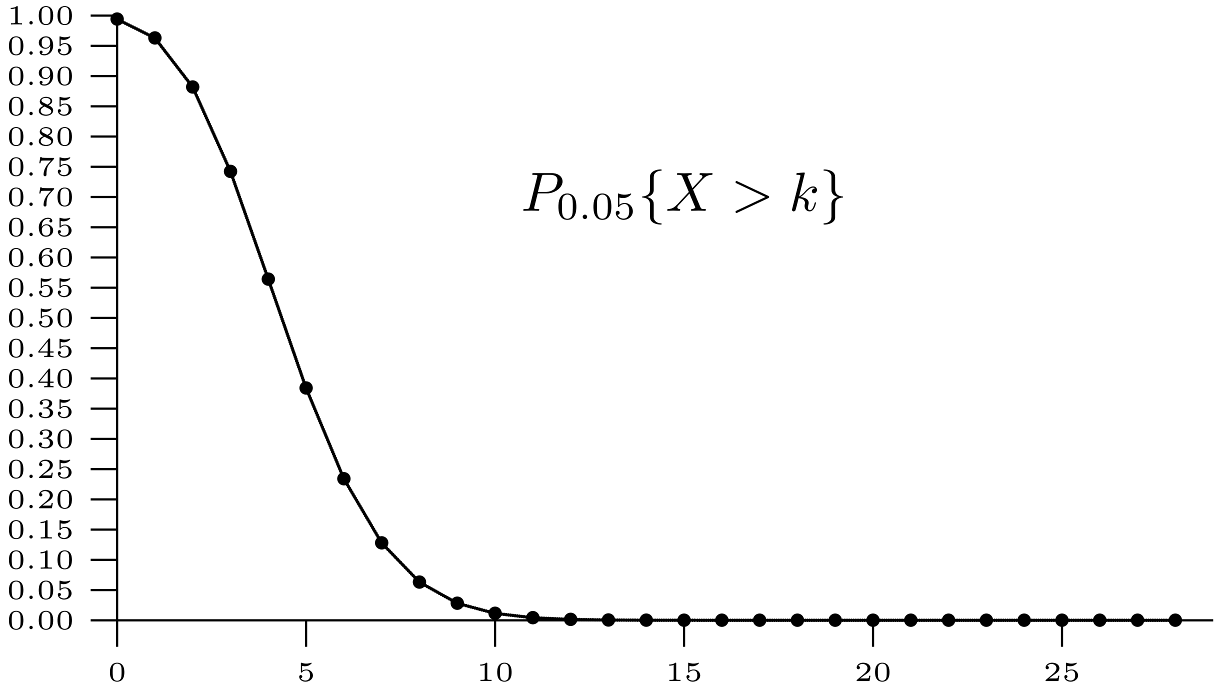
# Weryfikacja hipotez statystycznych

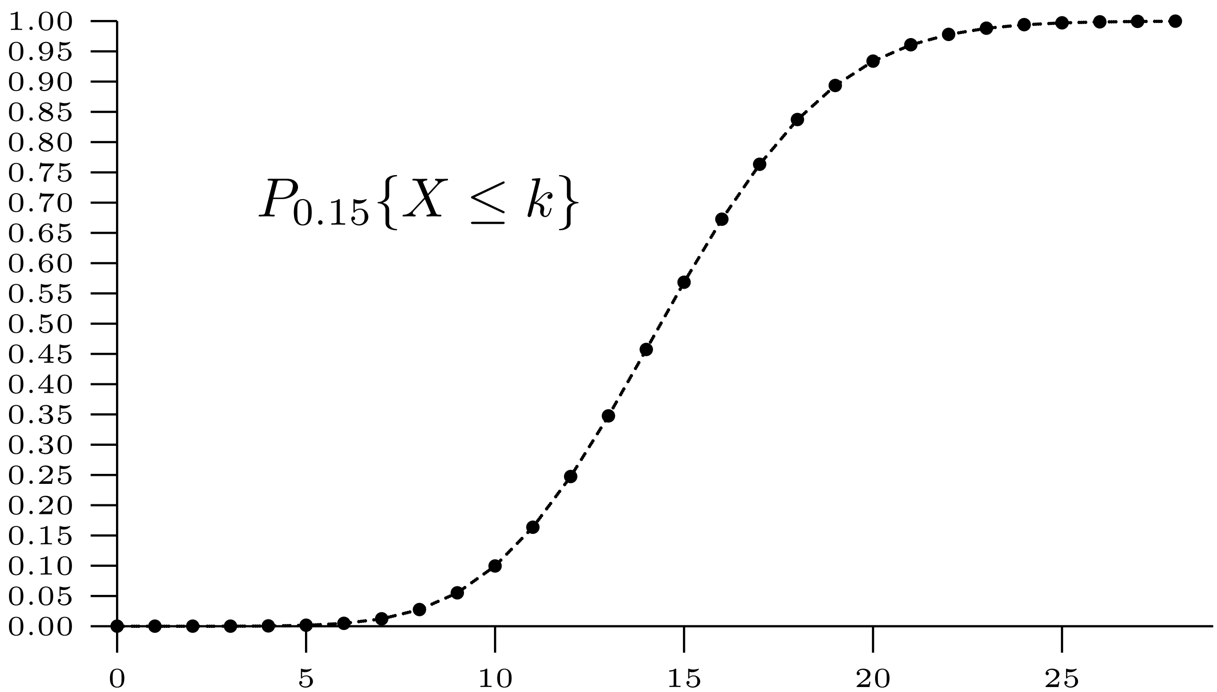
***Przykład (wstępny).*** Producent twierdzi, że wadliwość produkcji wynosi 5%. My podejrzewamy, że rzeczywista wadliwość produkcji wynosi 15%. Pobieramy próbę stuelementową i zliczamy ilość $X$ sztuk wadliwych. Model statystyczny

$$\{\{0, 1, \ldots, 100\}, \{B(100, \theta), \theta \in \Theta = \{0.05, 0.15\}\}\}$$



Jeżeli $X \le k$, to uznać $\theta = 0.05$
Jeżeli $X > k$, to uznać $\theta = 0.15$



Model statystyczny:

$$(\mathcal{X}, \{P_\theta,\ \theta \in \Theta\})$$

Hipoteza statystyczna: podzbiór $\Theta_0$ zbioru $\Theta$.

$\Theta_0$ nazywamy **hipotezą zerową**

$\Theta_1 = \Theta \setminus \Theta_0$ nazywamy **hipotezą alternatywną**

Jeżeli zbiór $\Theta_0$ jest jednoelementowy, to mówimy o hipotezie **prostej**, w przeciwnym przypadku mówimy o hipotezie **złożonej**.

Zapis klasyczny

$$H_0 : \theta \in \Theta_0 \qquad H_1 : \theta \in \Theta_1$$

**Test statystyczny**: procedura statystyczna, w wyniku której podejmujemy jedną z dwóch decyzji:
*odrzucić hipotezę zerową* $H_0$
lub
*nie ma podstaw do odrzucenia hipotezy zerowej* $H_0$.

Formalnie: test hipotezy $H_0$ utożsamiamy z funkcją $\phi : \mathcal{X} \to \{0, 1\}$

$$\phi(x) = \begin{cases} 1 & \text{odrzucić } H_0 \\ 0 & \text{nie odrzucać } H_0 \end{cases}$$

Test zrandomizowany: funkcja $\phi : \mathcal{X} \to [0, 1]$

$$\phi(x) \begin{cases} = 1 & \text{odrzucić } H_0 \\ \in (0,1) & \begin{array}{l}\text{na podstawie niezależnego od } X \\ \text{mechanizmu losowego odrzucić } H_0 \\ \text{z prawdopodobieństwem } \phi(x)\end{array} \\ = 0 & \text{nie odrzucać } H_0 \end{cases}$$

**Obszar krytyczny**: $\{x \in \mathcal{X} : \phi(x) = 1\}$

**Błąd I rodzaju**: błąd polegający na odrzuceniu hipotezy zerowej $H_0$, gdy w rzeczywistości jest ona prawdziwa.

**Poziom istotności**. Niech $\alpha \in (0, 1)$. Test $\phi$ jest na poziomie istotności $\alpha$, jeżeli

$$E_\theta \phi(X) = P_\theta\{\phi(X) = 1\} \leq \alpha, \ \forall \theta \in \Theta_0$$

**Rozmiar testu**:

$$\sup_{\theta \in \Theta_0} E_\theta \phi(X) = \sup_{\theta \in \Theta_0} P_\theta\{\phi(X) = 1\}$$

**Błąd II rodzaju**: błąd polegający na nieodrzuceniu hipotezy zerowej $H_0$, gdy w rzeczywistości jest ona fałszywa.

**Moc testu**:

$$\Theta_1 \in \theta \rightarrow E_\theta \phi(X) = P_\theta\{\phi(X) = 1\}$$

**Twierdzenie 10.1** *(lemat Neymana–Pearsona)* *Niech $P_{\theta_0}$ oraz $P_{\theta_1}$ będą rozkładami prawdopodobieństwa o gęstościach $f_0$ i $f_1$. Niech $\alpha \in (0,1)$ będzie ustaloną liczbą*

**a.** (Istnienie testu) *Istnieją takie stałe $t$ i $\gamma$, że*

$$\phi(x) = \begin{cases} 1, & \text{gdy } f_1(x) > tf_0(x) \\ \gamma, & \text{gdy } f_1(x) = tf_0(x) \\ 0, & \text{gdy } f_1(x) < tf_0(x) \end{cases}$$

*jest testem hipotezy $H_0 : \theta = \theta_0$ przeciwko $H_1 : \theta = \theta_1$ na poziomie istotności $\alpha$, tzn.*

$$(*) \qquad E_{\theta_0}\phi(X) = \alpha$$

**b.** (Dostateczność) *Jeżeli test $\phi$ spełnia warunek $(*)$ i dla pewnego $t$ warunek*

$$(**) \qquad \phi(x) = \begin{cases} 1, & \text{gdy } f_1(x) > tf_0(x) \\ 0, & \text{gdy } f_1(x) < tf_0(x) \end{cases}$$

*to $\phi$ jest testem najmocniejszym na poziomie istotności $\alpha$*

**c.** (Konieczność) *Jeżeli $\phi$ jest testem najmocniejszym na poziomie istotności $\alpha$, to spełnia on warunek $(**)$* ■

***Przykład (Model dwumianowy).***
Niech $\theta_0 < \theta_1 \in (0,1)$. Model statystyczny

$$\{\{0,1,\ldots,n\},\{B(n,\theta),\theta\in\Theta=\{\theta_0,\theta_1\}\}\}$$

$$H_0:\theta=\theta_0 \qquad H_1:\theta=\theta_1$$

$$f_0(x)=\binom{n}{x}\theta_0^x(1-\theta_0)^{n-x}$$

$$f_1(x)=\binom{n}{x}\theta_1^x(1-\theta_1)^{n-x}$$

Konstrukcja testu $\phi(x)$

$$\phi(x)=1\Leftrightarrow\frac{f_1(x)}{f_0(x)}>t\Leftrightarrow\left[\frac{\theta_1(1-\theta_0)}{\theta_0(1-\theta_1)}\right]^x>t\left[\frac{1-\theta_0}{1-\theta_1}\right]^n$$

Obszar krytyczny: $\{x>k\}$

Stała $k$ dobrana jest tak, że

$$E_{\theta_0}\phi(X)=P_{\theta_0}\{X>k\}\le\alpha$$

Niech $n = 100$, $\theta_0 = 0.05$, $\theta_1 = 0.15$, $\alpha = 0.05$

| $k$ | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|
| $P_{0.05}\{x>k\}$ | 0.12796 | 0.06309 | 0.02819 | 0.01147 | 0.00427 |
| $P_{0.05}\{x=k\}$ | 0.10603 | 0.06487 | 0.03490 | 0.01672 | 0.00720 |

Test niezrandomizowany

$$\phi(x) = \begin{cases} 1, & \text{jeżeli } x > 9, \\ 0, & \text{jeżeli } x \leq 9. \end{cases}$$

Rozmiar testu: $P_{0.05}\{x > 9\} = 0.02819$
Błąd II rodzaju: $P_{0.15}\{x \leq 9\} = 0.05509$

Test zrandomizowany

$$\phi(x) = \begin{cases} 1, & \text{jeżeli } x > 9, \\ 0.62495, & \text{jeżeli } x = 9 \\ 0, & \text{jeżeli } x < 9. \end{cases}$$

Rozmiar testu

$$\begin{aligned} & P_{0.05}\{x > 9\} + 0.62495 \cdot P_{0.05}\{x = 9\} \\ & \qquad = 0.02819 + 0.62495 \cdot 0.03490 = 0.05 \end{aligned}$$

. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

***Przykład .*** Niech $a, b > 1$. Model statystyczny

$$\{(0,1), \{U(0,1), B(a,b)\}\}$$

$$H_0 : U(0,1) \qquad H_1 : B(a,b)$$

$$f_0(x) = 1 \cdot \mathbf{1}_{(0,1)}(x)$$
$$f_1(x) \propto x^{a-1}(1-x)^{b-1}\mathbf{1}_{(0,1)}(x)$$

Konstrukcja testu $\phi(x)$

$$\phi(x) = 1 \Leftrightarrow \frac{f_1(x)}{f_0(x)} > t \Leftrightarrow x^{a-1}(1-x)^{b-1} > t$$

Obszar krytyczny: $\{x_0 < x < x_1\}$

Liczby $x_0$, $x_1$ dobrane są tak, że

$$E_{\theta_0}\phi(X) = P_{U(0,1)}\{x_0 < X < x_1\} \leq \alpha$$

$$x_0^{a-1}(1-x_0)^{b-1} = x_1^{a-1}(1-x_1)^{b-1}$$

Ponieważ $x_1 = x_0 + \alpha$, więc

$$\left[\frac{x_0+\alpha}{x_0}\right]^{a-1}\left[\frac{1-x_0-\alpha}{1-x_0}\right]^{b-1} = 1$$

| $a$ | $b$ | $x_0$ | $x_1$ |
|---|---|---|---|
| 2 | 2 | 0.47500 | 0.52500 |
| 2 | 3 | 0.30865 | 0.35865 |
| 2 | 4 | 0.22556 | 0.27556 |
| 3 | 2 | 0.64135 | 0.69135 |
| 3 | 3 | 0.47500 | 0.52500 |
| 3 | 4 | 0.37517 | 0.42517 |
| 4 | 2 | 0.72444 | 0.77444 |
| 4 | 3 | 0.57483 | 0.62483 |
| 4 | 4 | 0.47500 | 0.52500 |

. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

# Hipotezy złożone

Model statystyczny:

$$(\mathcal{X}, \{P_\theta,\ \theta \in \Theta\})$$

$$H_0 : \theta \in \Theta_0 \qquad H_1 : \theta \in \Theta_1$$

$H_0$ i/lub $H_1$ złożone

Iloraz wiarogodności

$$\lambda(x) = \frac{\sup\limits_{\theta \in \Theta_1} f_\theta(x)}{\sup\limits_{\theta \in \Theta_0} f_\theta(x)} \quad \text{lub} \quad \lambda(x) = \frac{\sup\limits_{\theta \in \Theta} f_\theta(x)}{\sup\limits_{\theta \in \Theta_0} f_\theta(x)}$$

Test

$$\phi(x) = \begin{cases} 1, & \text{gdy } \lambda(x) > t \\ \gamma, & \text{gdy } \lambda(x) = t \\ 0, & \text{gdy } \lambda(x) < t \end{cases}$$

Dobór stałej $t$

$$\sup_{\theta \in \Theta_0} E_\theta \phi(X) \le \alpha$$

***Przykład (Model dumianowy).*** Niech $\theta_0 \in (0,1)$. Model statystyczny

$$\{\{0,1,\ldots,n\},\{B(n,\theta),\theta\in\Theta=(0,1)\}\}$$

$$H_0:\theta\le\theta_0 \qquad H_1:\theta>\theta_0$$

$$f_\theta(x)=\binom{n}{x}\theta^x(1-\theta)^{n-x}$$

Ponieważ

$$\sup_{\theta\in(0,1)} f_\theta(x)=f_{\frac{x}{n}}(x)=\binom{n}{x}\left(\frac{x}{n}\right)^x\left(1-\frac{x}{n}\right)^{n-x}$$

więc

$$\sup_{\theta\in\Theta_0} f_\theta(x)=\begin{cases}\binom{n}{x}\left(\frac{x}{n}\right)^x\left(1-\frac{x}{n}\right)^{n-x}, & \text{jeżeli } \frac{x}{n}\le\theta_0,\\ \binom{n}{x}\theta_0^x(1-\theta_0)^{n-x}, & \text{jeżeli } \frac{x}{n}>\theta_0,\end{cases}$$

oraz

$$\sup_{\theta\in\Theta_1} f_\theta(x)=\begin{cases}\binom{n}{x}\theta_0^x(1-\theta_0)^{n-x}, & \text{jeżeli } \frac{x}{n}\le\theta_0,\\ \binom{n}{x}\left(\frac{x}{n}\right)^x\left(1-\frac{x}{n}\right)^{n-x}, & \text{jeżeli } \frac{x}{n}>\theta_0.\end{cases}$$

Zatem

$$\lambda(x) = \frac{\sup\limits_{\theta\in\Theta_1} f_\theta(x)}{\sup\limits_{\theta\in\Theta_0} f_\theta(x)}$$

$$= \begin{cases} \dfrac{\binom{n}{x}\theta_0^x(1-\theta_0)^{n-x}}{\binom{n}{x}\left(\frac{x}{n}\right)^x\left(1-\frac{x}{n}\right)^{n-x}}, & \text{jeżeli } \frac{x}{n} \le \theta_0, \\[2ex] \dfrac{\binom{n}{x}\left(\frac{x}{n}\right)^x\left(1-\frac{x}{n}\right)^{n-x}}{\binom{n}{x}\theta_0^x(1-\theta_0)^{n-x}}, & \text{jeżeli } \frac{x}{n} > \theta_0. \end{cases}$$

$\lambda(x)$ jest rosnąca ze względu na $x$

$$\lambda(x) > t \Leftrightarrow x > k$$

Dobór stałej $k$

$$\sup_{\theta\le\theta_0} P_\theta\{X > k\} \overset{?}{=} P_{\theta_0}\{X > k\} \le \alpha$$

. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

***Przykład (Model gaussowski).*** Niech $\mu_0 \in \mathbb{R}$. Model dla próby $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$:

$$\left(\mathbb{R}^n, \left\{N\left(\mu\mathbf{1}_n, \sigma^2\boldsymbol{I}_n\right),\ \theta = (\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}_+\right\}\right)$$

$$\Theta = \mathbb{R}\times\mathbb{R}_+ \quad \Theta_0 = \{\mu_0\}\times\mathbb{R}_+ \quad \Theta_1 = (\mathbb{R}\setminus\{\mu_0\})\times\mathbb{R}_+$$

$$H_0 : \mu = \mu_0 \qquad H_1 : \mu \neq \mu_0$$

$$f_{\mu,\sigma}(\boldsymbol{x}) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left\{-\frac{1}{2}\sum_{i=1}^n \left(\frac{x_i - \mu}{\sigma}\right)^2\right\}$$

Niech

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})^2 \quad \tilde{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \mu_0)^2$$

$$\sup_{\theta\in\Theta_0} f_{\mu,\sigma}(x) \overset{?}{=} f_{\mu_0,\tilde{\sigma}}(x) \overset{?}{=} \left(\frac{1}{\tilde{\sigma}\sqrt{2\pi}}\right)^n \exp\left\{-\frac{n}{2}\right\}$$

$$\sup_{\theta\in\Theta} f_{\mu,\sigma}(x) \overset{?}{=} f_{\bar{x},\hat{\sigma}}(x) \overset{?}{=} \left(\frac{1}{\hat{\sigma}\sqrt{2\pi}}\right)^n \exp\left\{-\frac{n}{2}\right\}$$

Iloraz wiarogodności

$$\lambda(x) = \frac{\sup\limits_{\theta\in\Theta} f_{\mu,\sigma}(x)}{\sup\limits_{\theta\in\Theta_0} f_{\mu,\sigma}(x)} = \left(\frac{\tilde{\sigma}}{\hat{\sigma}}\right)^n$$

Ponieważ

$$\sum_{i=1}^n (X_i - \mu_0)^2 = \sum_{i=1}^n (X_i - \bar{X} + \bar{X} - \mu_0)^2$$

$$\overset{?}{=} \sum_{i=1}^n (X_i - \bar{X})^2 + n(\bar{X} - \mu_0)^2$$

więc

$$\lambda(x) = \left(1 + \frac{n(\bar{X} - \mu_0)^2}{\sum_{i=1}^n (X_i - \bar{X})^2}\right)^{\frac{n}{2}}$$

Test

$$\phi(x) = 1 \Leftrightarrow \lambda(x) > t \Leftrightarrow \frac{n(\bar{X} - \mu_0)^2}{\sum_{i=1}^n (X_i - \bar{X})^2} > t'$$

Dobór stałej $t'$

$$E_{H_0}\phi(X) = \alpha$$

**a.** jeżeli $H_0$ jest prawdziwa, to

$$\sqrt{n}(\bar{X} - \mu_0) \sim N(0, \sigma^2) \quad \text{czyli} \quad \frac{n(\bar{X} - \mu_0)^2}{\sigma^2} \sim \chi_1^2$$

**b.** dla wszystkich $\theta \in \Theta$

$$\frac{1}{\sigma^2} \sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2 \sim \chi_{n-1}^2$$

**c.** $\bar{X}$ oraz $\sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2$ są niezależne
Jeżeli hipoteza $H_0$ jest prawdziwa, to

$$\frac{n(\bar{X} - \mu_0)^2}{\frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2} \sim F_{1,n-1}$$

$H_0$ jest odrzucana na poziomie istotności $\alpha$, jeżeli

$$(*) \qquad \frac{n(\bar{X} - \mu_0)^2}{\frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2} > F\left(\alpha; 1, n-1\right)$$

Ponieważ $t_v = \sqrt{F_{1,v}}$, więc $(*)$ jest równoważne

$$(\clubsuit) \qquad \frac{|\bar{X} - \mu_0|}{S} \sqrt{n} > t(\alpha; n-1)$$

Test $(\clubsuit)$ nazywa się **testem Studenta**

. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

***Przykład (Model gaussowski).*** Wariancja $\sigma^2$ jest znana. Niech $\mu_0 \in \mathbb{R}$.

$$\left(\mathbb{R}^n, \left\{N\left(\mu \mathbf{1}_n, \sigma^2 \boldsymbol{I}_n\right),\ \theta = \mu \in \mathbb{R}\right\}\right)$$

$$\Theta = \mathbb{R} \quad \Theta_0 = (-\infty, \mu_0\rangle \quad \Theta_1 = (\mu_0, \infty)$$

$$H_0 : \mu \leq \mu_0 \qquad H_1 : \mu > \mu_0$$

Stosując technikę poprzedniego przykładu otrzymujemy test o obszarze krytycznym

$$\frac{\bar{X} - \mu_0}{\sigma}\sqrt{n} > u_{1-\alpha}$$

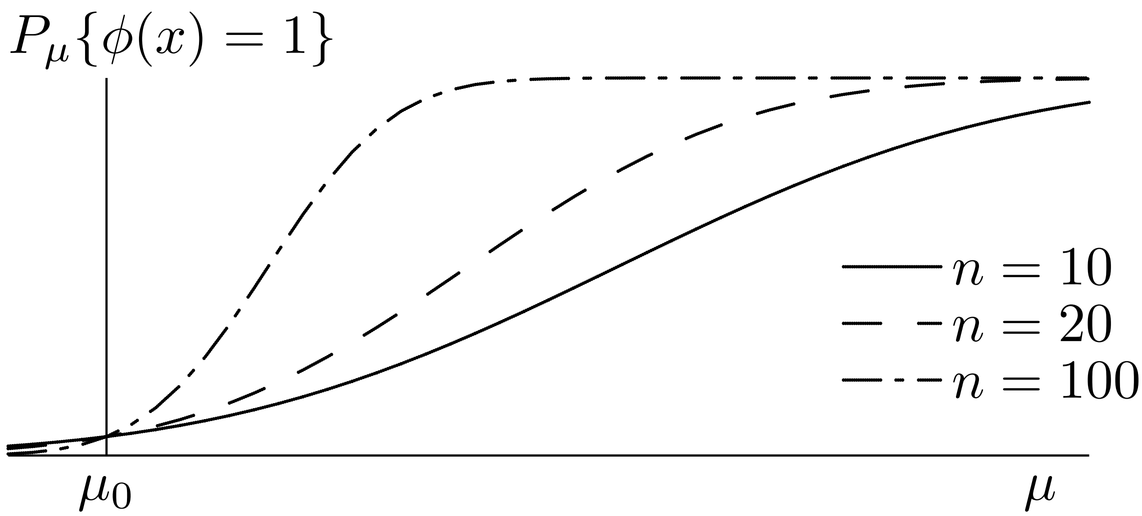
**Moc testu.** Niech $\mu > \mu_0$. Prawdopodobieństwo odrzucenia $H_0$

$$P_\mu\left\{\frac{\bar{X} - \mu_0}{\sigma}\sqrt{n} > u_{1-\alpha}\right\} =$$

$$P_\mu\left\{\frac{\bar{X} - \mu}{\sigma}\sqrt{n} > u_\alpha - \frac{\mu - \mu_0}{\sigma}\sqrt{n}\right\} =$$

$$1 - \Phi\left(u_\alpha - \frac{\mu - \mu_0}{\sigma}\sqrt{n}\right)$$

Moc testu jest zależna od $(\mu - \mu_0)/\sigma$




**Liczność próby.** Niech $(\mu - \mu_0)/\sigma = x_0$ będzie daną liczbą. Powiedzmy, że interesuje nas osiągnięcie dla tej wartości mocy co najmniej $\gamma$. Szukamy takiego $n$, że

$$1 - \Phi(u_{1-\alpha} - x_0\sqrt{n}) \geq \gamma$$

Rozwiązanie:

$$u_{1-\alpha} - x_0\sqrt{n} \leq u_{1-\gamma}$$

Stąd

$$n \geq \left[\frac{u_{1-\alpha} - u_{1-\gamma}}{x_0}\right]^2$$

| | $\gamma$ ($\alpha$=0.05) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| $x_0$ | 0.60 | 0.65 | 0.70 | 0.75 | 0.80 | 0.85 | 0.90 | 0.95 |
| 0.01 | 48987 | 55003 | 61720 | 69403 | 78488 | 89783 | 105073 | 129946 |
| 0.05 | 1959 | 2200 | 2468 | 2776 | 3139 | 3591 | 4202 | 5197 |
| 0.10 | 489 | 550 | 617 | 694 | 784 | 897 | 1050 | 1299 |
| 0.15 | 217 | 244 | 274 | 308 | 348 | 399 | 466 | 577 |
| 0.20 | 122 | 137 | 154 | 173 | 196 | 224 | 262 | 324 |
| 0.25 | 78 | 88 | 98 | 111 | 125 | 143 | 168 | 207 |
| 0.30 | 54 | 61 | 68 | 77 | 87 | 99 | 116 | 144 |
| 0.35 | 39 | 44 | 50 | 56 | 64 | 73 | 85 | 106 |
| 0.40 | 30 | 34 | 38 | 43 | 49 | 56 | 65 | 81 |
| 0.45 | 24 | 27 | 30 | 34 | 38 | 44 | 51 | 64 |
| 0.50 | 19 | 22 | 24 | 27 | 31 | 35 | 42 | 51 |
| 0.55 | 16 | 18 | 20 | 22 | 25 | 29 | 34 | 42 |
| 0.60 | 13 | 15 | 17 | 19 | 21 | 24 | 29 | 36 |
| 0.65 | 11 | 13 | 14 | 16 | 18 | 21 | 24 | 30 |
| 0.70 | 9 | 11 | 12 | 14 | 16 | 18 | 21 | 26 |
| 0.75 | 8 | 9 | 10 | 12 | 13 | 15 | 18 | 23 |
| 0.80 | 7 | 8 | 9 | 10 | 12 | 14 | 16 | 20 |
| 0.85 | 6 | 7 | 8 | 9 | 10 | 12 | 14 | 17 |
| 0.90 | 6 | 6 | 7 | 8 | 9 | 11 | 12 | 16 |
| 0.95 | 5 | 6 | 6 | 7 | 8 | 9 | 11 | 14 |
| 1.00 | 4 | 5 | 6 | 6 | 7 | 8 | 10 | 12 |

. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

# Wartość p–value

***Przykład (Model gaussowski).*** Wariancja $\sigma^2$ jest znana. Niech $\mu_0 \in \mathbb{R}$.

$$\left(\mathbb{R}^n, \left\{N\left(\mu\mathbf{1}_n, \sigma^2\boldsymbol{I}_n\right),\ \theta=\mu\in\mathbb{R}\right\}\right)$$

$$\Theta=\mathbb{R} \quad \Theta_0=\{\mu_0\} \quad \Theta_1=\mathbb{R}\setminus\{\mu_0\}$$

$$H_0 : \mu=\mu_0 \qquad H_1 : \mu\neq\mu_0$$

Test oparty na statystyce

$$Z(X)=n\left(\frac{\bar{X}-\mu_0}{\sigma}\right)^2$$

Duże wartości $Z$ przemawiają przeciwko $H_0$
Jeżeli $H_0$ jest prawdziwa, to $Z(X)$ ma rozkład chi–kwadrat z jednym stopniem swobody.
Niech $Z(x)$ będzie wartością statystyki $Z$ zaobserwowaną w próbie. Liczbę

$$p=P\{\chi_1^2>Z(x)\}$$

nazywamy $p$–value

. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .