# 4. Hypothesis tests

Estimation: confidence intervals

Now: testing hypotheses against another

typical set-up:

$H_0$ *null hypothesis*

$H_1$ *alternative hypothesis*

**Occam's razor:** accept the simplest explanation unless evidence against it

$H_0$ typically is the simplest explanation

Thus testing is not symmetric: we would not like to reject $H_0$ unless the evidence against it is too strong

*Paradox:* Often one wants to establish a more complex explanation, hence seeks to reject $H_0$

*Examples:*

1. Toss a coin, either fair or biased .7 towards head

   $p = \mathbf{P}(Head)$

   $H_0 : p = \frac{1}{2}$

   $H_1 : p = 0.7$

   (symmetric problem)

2. Gender discrimination on jury selection?

   $p = \mathbf{P}(female)$

   $H_0 : p = \frac{1}{2}$: no discrimination

   $H_1 : p < \frac{1}{2}$

3. *Swain vs. Alabama*, see

   `http://www.stat.ucla.edu/cases/swain/`

*Definition.* A *statistical hypothesis* specifies a family of distributions of the observations

In our framework: a set of parameters; if $\theta \in \Theta$ is the whole parameter space, we would have

$H_0 : \theta \in \Theta_0$

$H_1 : \theta \in \Theta_1$

where $\Theta_0 \subset \Theta, \Theta_1 \subset \Theta$, and $\Theta_0 \cap \Theta_1 = \emptyset$

The hypothesis is *simple* if there is only one parameter value in the set, i.e. if it specifies the distribution completely; otherwise it is called *composite*

**Test:** Observe $x_1, \ldots, x_n$, calculate test statistic $t(x_1, \ldots, x_n)$; reject $H_0$ if $t$ takes on a value that would be very unusual if $H_0$ was true

A test is equivalent to specifying a set $C$ of sample values such that:

If $\mathbf{x} \in C$ then we reject $H_0$ in favour of $H_1$

This set $C$ is called the *critical region* of the test.

*Example* Flip coin $n$ times, let $x_i = 1$ if the coin lands head in the ith toss, and $x_i = 0$ if in the ith toss the coin lands tail

*Assume* that the coin flips are realizations of i.i.d. Bernoulli random variables with unknown probability $p$ of success

Put

$$T = t(\mathbf{X}) = \sum_{i=1}^{n} X_i \sim Bin(n, p)$$

$H_0 : p = \frac{1}{2}$

$H_1 : p = 0.7$

Reject $H_0$ in favour of $H_1$ if $T$ is large: $T > c$

Specify *level* $\alpha$: Want $c$ such that

$$\mathbf{P}_0(T > c) \leq \alpha$$

where $\mathbf{P}_0$ denotes the probability under $H_0$, i.e. here $Bin(n, \frac{1}{2})$

So want $c$ such that

$$\sum_{k>c} \binom{n}{k} 2^{-n} = \alpha$$

Find $c$ from Table; e.g. $n = 15, \alpha = 0.05$ : $\mathbf{P}_0(T > 11) = 0.018, \mathbf{P}_0(T > 10) = 0.059$, so choose $c = 11$

Note: Under $H_1$, $T \sim Bin(15, .7)$ and $\mathbf{P}_1(T > 11) = 0.297$ is not very large neither

*Notation:* $\mathbf{P}_\theta$ is the probability if $\theta$ is the true parameter

$H_0 : \theta \in \Theta_0$

$H_1 : \theta \in \Theta_1$

The *size* of the test is

$$\sup_{\theta \in \Theta_0} \mathbf{P}_\theta(C),$$

it is also called the *Type I error probability*

If the size of the test is $\alpha$ and if $\mathbf{x} \in C$, then we say that the test is *significant* at the $100(1-\alpha)\%$ level

The *power* of the test is given by the curve

$$\mathbf{P}_\theta(C), \theta \notin \Theta_0$$

For $H_1$ simple, the *Type II error* is $1 - power$

*Example*

$x_1, \ldots, x_n$ random sample from $\mathcal{N}(\mu, \sigma^2)$, $\sigma^2$ known

$H_0; \mu = \mu_0$

$H_1 : \mu \neq \mu_0$

$$t(\mathbf{x}) = \frac{\bar{x} - \mu_0}{\sqrt{\sigma^2/n}}$$

then

$$1-\alpha = \mathbf{P}_0(\bar{X} - z_{\alpha/2}\sqrt{\sigma^2/n} < \mu_0 < \bar{X} + z_{\alpha/2}\sqrt{\sigma^2/n})$$

so for a test of size $\alpha$ we can choose

$$\begin{aligned} C &= \{\mathbf{x} : \bar{x} > \mu_0 + z_{\alpha/2}\sqrt{\sigma^2/n}\} \\ &\quad \cup \{\mathbf{x} : \bar{x} < \mu_0 - z_{\alpha/2}\sqrt{\sigma^2/n}, \infty\} \end{aligned}$$

*Power:* If $\mu_1 \neq \mu_0$,

$$
\begin{aligned}
&\mathbf{P}_{\mu_1}(C) \\
&= \mathbf{P}_{\mu_1}(\bar{X} - z_{\alpha/2}\sqrt{\sigma^2/n} > \mu_0) \\
&\quad + \mathbf{P}_{\mu_1}(\mu_0 > \bar{X} + z_{\alpha/2}\sqrt{\sigma^2/n}) \\
&= \mathbf{P}_{\mu_1}(\bar{X} - \mu_1 > z_{\alpha/2}\sqrt{\sigma^2/n} + (\mu_0 - \mu_1)) \\
&\quad + \mathbf{P}_{\mu_1}(\bar{X} - \mu_1 < -z_{\alpha/2}\sqrt{\sigma^2/n} + (\mu_0 - \mu_1)) \\
&= 1 - \Phi\left(z_{\alpha/2} + \frac{\mu_0 - \mu_1}{\sqrt{\sigma^2/n}}\right) \\
&\quad + \Phi\left(-z_{\alpha/2} + \frac{\mu_0 - \mu_1}{\sqrt{\sigma^2/n}}\right)
\end{aligned}
$$

increases with $|\mu_0 - \mu_1|$

Can adjust $n$ such that the test has a specific power

**Connection with confidence intervals:**

A pivot can also be used as a test statistic

confidence interval = the set of parameters which we do not reject

*In previous example:*

$$\begin{aligned} C &= \{\mathbf{x} : \bar{x} > \mu_0 + z_{\alpha/2}\sqrt{\sigma^2/n}\} \\ &\quad \cup \{\mathbf{x} : \bar{x} < \mu_0 - z_{\alpha/2}\sqrt{\sigma^2/n}, \infty\} \end{aligned}$$

So accept if $\mathbf{x} \in C^c$, i.e. if

$$\mu_0 - z_{\alpha/2}\sqrt{\sigma^2/n} < \bar{x} < \mu_0 + z_{\alpha/2}\sqrt{\sigma^2/n}$$

i.e. accept for all $\mu_0$ for which

$$\bar{x} - z_{\alpha/2}\sqrt{\sigma^2/n} < \mu_0 < \bar{x} + z_{\alpha/2}\sqrt{\sigma^2/n}$$

## Likelihood ratio tests

*Example:* Toss a coin $n$ times, either fair or biased .7 towards head

$p = \mathbf{P}(Head) = \mathbf{P}(X_i = 1)$

$H_0 : p = \frac{1}{2}$

$H_1 : p = 0.7$

use *likelihood ratio (LR)*

$$\frac{L(\frac{1}{2}; \mathbf{x})}{L(0.7; \mathbf{x})}$$

reject if LR small

Here: $t(\mathbf{x}) = \sum_{i=1}^{n} x_i$

$$\begin{aligned} \frac{L(\frac{1}{2}; \mathbf{x})}{L(0.7; \mathbf{x})} &= \frac{\binom{n}{t} 2^{-n}}{\binom{n}{t} (0.7)^t (0.3)^{n-t}} \\ &= \left(\frac{1}{2}\right)^n \left(\frac{10}{3}\right)^n \left(\frac{3}{7}\right)^t \end{aligned}$$

reject if LR small: if $t$ is large

*General*: $H_0, H_1$ simple

$H_0 : \theta = \theta_0$

$H_1 : \theta = \theta_1$

The *likelihood ratio (LR)* is

$$\lambda(\mathbf{x}) := \frac{L(\theta_0; \mathbf{x})}{L(\theta_1; \mathbf{x})}$$

The *likelihood ratio test (LRT)* rejects $H_0$ when $\lambda(\mathbf{x})$ small

Or: give the *p-value*

$$p = \mathbf{P}(\lambda(\mathbf{X}) < \lambda(\mathbf{x}) | \theta = \theta_0)$$

reject if $p$-value small

$p-value$ is the probability to observe a LR this small or smaller when $H_0$ is true

*Example*

$\mathcal{N}(\mu, \sigma^2)$, with $\sigma^2$ known

$H_0 : \mu = \mu_0$

$H_1 : \mu = \mu_1$, where $\mu_1$ is some fixed number;

assume $\mu_1 < \mu_0$

Fix level $\alpha$

$$\lambda(\mathbf{x}) = \frac{(2\pi\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i - \mu_0)^2}}{(2\pi\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i - \mu_1)^2}}$$

LR small $\iff$ log LR small:

$$\begin{aligned} \log LR &= -\frac{1}{2\sigma^2} \sum_{i=1}^n \{(x_i - \mu_0)^2 - (x_i - \mu_1)^2\} \\ &= -\frac{1}{2\sigma^2} \{2n\bar{x}(\mu_1 - \mu_0) + n(\mu_0^2 - \mu_1^2)\} \end{aligned}$$

as $\mu_1 < \mu_0$ : $\log LR$ increases $\iff$ $\bar{x}$ increases

so: reject $H_0$ if $\bar{x}$ small

$\alpha \leq \mathbf{P}_0(\bar{X} < k)$

so: reject $H_0$ iff

$$\bar{X} < \mu_0 - z_\alpha \frac{\sigma}{\sqrt{n}}$$

*Power:*

$$1 - \Phi\left(\frac{\mu_0 - \mu_1}{\frac{\sigma}{\sqrt{n}}} - z_\alpha\right) \geq 1 - \Phi(-z_\alpha) = \alpha$$

increases if $\mu_0 - \mu_1$ increases

is a *one-sided test*

Sample-size calculation:

If we wanted to be near-certain to reject $H_0$ when $\mu = \mu_0 - \delta$, say, and have size 0.05, then we could fix $n$ to force $power(\mu) = 0.99$ at $\mu = \mu_0 - \delta$: i.e.,

$$0.99 = 1 - \Phi(\delta\sqrt{n}/\sigma - 1.645)$$

This equation reduces to $2.326 = \delta\sqrt{n}/\sigma - 1.645$, so that

$$n = \sigma^2(1.645 + 2.326)^2/\delta^2$$

is the required sample size

## Constructing good tests: The Neyman-Pearson Lemma

*Neyman-Pearson Lemma:* Suppose the LRT that rejects $H_0$ when $\lambda(\mathbf{x}) < c$ has significance level $\alpha$. Then any other test which has significance level $\alpha^* \le \alpha$ has power less or equal to that of the LRT.

I.e.: The LRT is most powerful among all tests of level $\alpha$

*Proof*

Let $C = (-\infty, A)$ be the critical region of the LRT, power $\beta$

Pick any other test of level $\alpha^* \leq \alpha$; let $C^*$ critical region, $\beta^*$ its power

then

$$
\begin{aligned}
\beta^* - \beta \\
&= P_1(\mathbf{X} \in C^*) - P_1(\mathbf{X} \in C) \\
&= \int_{C^*} L(\theta_1, \mathbf{x}) d\mathbf{x} - \int_C L(\theta_1, \mathbf{x}) d\mathbf{x} \\
&= \int_{C^* \cap C} L(\theta_1, \mathbf{x}) d\mathbf{x} + \int_{C^* \cap C^c} L(\theta_1, \mathbf{x}) d\mathbf{x} \\
&\quad - \int_{C \cap C^*} L(\theta_1, \mathbf{x}) d\mathbf{x} - \int_{C \cap (C^*)^c} L(\theta_1, \mathbf{x}) d\mathbf{x} \\
&= \int_{C^* \cap C^c} L(\theta_1, \mathbf{x}) d\mathbf{x} - \int_{C \cap (C^*)^c} L(\theta_1, \mathbf{x}) d\mathbf{x}.
\end{aligned}
$$

On $C$: $L(\theta_0, \mathbf{x}) < AL(\theta_1, \mathbf{x})$

On $C^c$: $L(\theta_0, \mathbf{x}) \geq AL(\theta_1, \mathbf{x})$,

for some $A > 0$. So

$$\begin{aligned} & A(\beta^* - \beta) \\ & \leq \int_{C^* \cap C^c} L(\theta_0, \mathbf{x}) d\mathbf{x} - \int_{C \cap (C^*)^c} L(\theta_0, \mathbf{x}) d\mathbf{x} \\ & = \int_{C^*} L(\theta_0, \mathbf{x}) d\mathbf{x} - \int_{C} L(\theta_0, \mathbf{x}) d\mathbf{x} \end{aligned}$$

(add and subtract $\int_{C \cap C^*} L(\theta_0, \mathbf{x}) d\mathbf{x}$) and so

$$A(\beta^* - \beta) \leq \alpha - \alpha^* \leq 0.$$

This finishes the proof.

*Back to example:* $\mathcal{N}(\mu, \sigma^2)$; $\sigma^2$ known

$H_0 : \mu = \mu_0$; $H_1 : \mu = \mu_1$, where $\mu_1 < \mu_0$

In LRT: reject $H_0$ iff

$$\bar{X} < \mu_0 - z_\alpha \frac{\sigma}{\sqrt{n}}$$

we only used that $\mu_1 < \mu_0$;

Neyman-Pearson Lemma gives: test is most powerful for all $\mu_1 < \mu_0$:

Test is *uniformly most powerful (UMP)*:

Let $H_1$ be a composite hypothesis. A test that is most powerful for *every* simple alternative in $H_1$ is said to be *uniformly most powerful*

Similarly: $\mathcal{N}(\mu, \sigma^2)$; $\sigma^2$ known

$H_0 : \mu = \mu_0$; $H_1 : \mu = \mu_1$, where $\mu_1 > \mu_0$

LRT: reject $H_0$ iff

$$\bar{X} > \mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}}$$

so test is UMP for $H_0 : \mu = \mu_0$; $H_1 : \mu > \mu_0$

But $H_0 : \mu = \mu_0$; $H_1 : \mu \neq \mu_0$: there is no UMP test

instead: reject if $|\bar{x} - \mu_0| > z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$

Quite often the case: tests are constructed which are good, but not optimal power on both sides of $\mu_0$: *two-sided tests*

**Constructing good tests: composite hypotheses**

*Example:* $\mathcal{N}(\mu, \sigma^2)$; test for $H_0 : \mu = \mu_0$; but $\sigma^2$ unknown:

is a *nuisance parameter*

*General:* $\theta = (\theta_1, \ldots, \theta_q)$, each $\theta_i \in \mathbf{R}$; assume:

take values in some open interval of $\mathbf{R}$

$\Theta$ parameter set

$H_0 : \theta \in \Theta_0 = \{\theta : \theta_1 = \theta_1^0, \ldots, \theta_r = \theta_r^0\}$

first $r$ parameter fixed, rest unconstrained

*Definition:* The *generalized likelihood ratio* is

$$\lambda(\mathbf{x}) = \frac{\sup_{\theta \in \Theta_0} L(\theta, \mathbf{x})}{\sup_{\theta \in \Theta} L(\theta, \mathbf{x})}$$

The expression $\sup_{\theta \in \Theta_0} L(\theta, \mathbf{x})$ is also called *profile likelihood*

The *generalized likelihood ratio test* rejects $H_0$ when $\lambda(\mathbf{x})$ is small

*Note:* Always $0 \leq \lambda(\mathbf{x}) \leq 1$ for the *generalized likelihood ratio*

*Example*

random sample from $\mathcal{N}(\mu, \sigma^2)$; $\sigma^2$ unknown, test

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

$\Theta_0 = \{\mu_0\} \times (0, \infty)$; $dim\Theta_0 = 1$

$\Theta = \mathbf{R} \times (0, \infty)$, $dim\Theta = 2$

$\theta = (\mu, \sigma^2)$

$$L(\theta, \mathbf{x}) = (2\pi\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i - \mu)^2}$$

For $\sup_{\theta\in\Theta_0} L(\theta, \mathbf{x})$: Under $H_0$, the likelihood is maximized for the m.l.e.

$$\hat{\sigma}_0^2 = \frac{1}{n}\sum_{i=1}^n (x_i - \mu_0)^2$$

In $\Theta$, the likelihood is maximized for the m.l.e.s

$$\hat{\mu} = \bar{x}, \quad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n (x_i - \bar{x})^2$$

Note that

$$
\begin{aligned}
L(\mu_0, \hat{\sigma}_0^2), \mathbf{x}) &= (2\pi\hat{\sigma}_0^2)^{-n/2} e^{-\frac{1}{2\hat{\sigma}_0^2}\sum_{i=1}^n (x_i-\mu_0)^2} \\
&= (2\pi\hat{\sigma}_0^2)^{-n/2} e^{-\frac{n}{2}}
\end{aligned}
$$

and similarly

$$
\begin{aligned}
L((\hat{\mu}, \hat{\sigma}^2), \mathbf{x}) &= (2\pi\hat{\sigma}^2)^{-n/2} e^{-\frac{1}{2\hat{\sigma}^2}\sum_{i=1}^n (x_i-\bar{x})^2} \\
&= (2\pi\hat{\sigma}^2)^{-n/2} e^{-\frac{n}{2}}
\end{aligned}
$$

So

$$
\begin{aligned}
\lambda(\mathbf{x}) &= \left\{ \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{\sum_{i=1}^n (x_i - \mu_0)^2} \right\}^{\frac{n}{2}} \\
&= \left\{ \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2 + n(\bar{x} - \mu_0)^2} \right\}^{\frac{n}{2}} \\
&= \left\{ 1 + \frac{(\bar{x} - \mu_0)^2}{1/n \sum_{i=1}^n (x_i - \bar{x})^2} \right\}^{-\frac{n}{2}}
\end{aligned}
$$

is monotone in $t^2(\mathbf{x})$, with $t$-statistic

$$
t(\mathbf{x}) = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}
$$

Small if $t^2$ large, i.e. if $t < 0$ and small, or if $t > 0$ and large

Testing $H_0 : \mu = \mu_0$ using the two-sided $t$-test is equivalent to testing it with the generalized LRT

Moreover, expand

$$
\begin{aligned}
&-2 \log \lambda(\mathbf{x}) \\
&\quad = \; -2\left(-\frac{n}{2}\right) \log\left\{1 + \frac{(\bar{x} - \mu_0)^2}{1/n \sum_{i=1}^n (x_i - \bar{x})^2}\right\} \\
&\quad \approx \; n \frac{(\bar{x} - \mu_0)^2}{1/n \sum_{i=1}^n (x_i - \bar{x})^2} \\
&\quad = \; \left\{\frac{\sqrt{n}(\bar{x} - \mu_0)^2}{\sqrt{1/n \sum_{i=1}^n (x_i - \bar{x})^2}}\right\}^2 \\
&\quad \approx \; (t(\mathbf{x}))^2 \\
&\quad \to \; \chi_1^2 \text{ in distribution}
\end{aligned}
$$

**Wilks' Theorem:** Under a few conditions, as $n \to \infty$, for the generalized LR $\lambda(\mathbf{x})$ we have, in distribution,

$$-2 \log \lambda(\mathbf{x}) \approx \chi^2_r,$$

where $r = dim\Theta - \dim \Theta_0$

So: reject $H_0$ at 5 % level, e.g., if the observed value of $-2 \log \lambda$ was greater than $\chi^2_r(.95)$

In large samples, $-2 \log \lambda$ is an approximate pivot, and can be used to obtain an approximate confidence set

$$\{(\theta^0_1, \ldots, \theta^0_r) : -2 \log \lambda(\mathbf{x}) < \chi^2_r(1 - \alpha)\}$$

When $r = 1$, and $n$ large, the generalized LRT is equivalent to constructing a two-sided test based on the MLE; test of size $\alpha$ rejects $H_0$ when

$$|\widehat{\theta}_1 - \theta_1^0|\sqrt{nI(\widehat{\theta})} > z_{\alpha/2}$$

(compare to confidence intervals)

random sample from $\mathcal{N}(\mu, \sigma^2)$; $\mu$ known, $\sigma^2$ unknown, test $H_0 : \sigma^2 = \sigma_0^2$ against $H_1 : \sigma^2 \neq \sigma_0^2$

$\Theta_0 = \{\sigma_0^2\}$

$\Theta = \{0, \infty\}$

$\theta = \sigma^2$

$$L(\theta, \mathbf{x}) = (2\pi\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i - \mu)^2}$$

In $\Theta$, the likelihood is maximized for the m.l.e.s

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n (x_i - \mu)^2$$

Note that

$$L(\sigma_0^2), \mathbf{x}) \;=\; (2\pi\sigma_0^2)^{-n/2} e^{-\frac{1}{2\sigma_0^2}\sum_{i=1}^n (x_i - \mu)^2}$$

and

$$L(\hat{\sigma}^2), \mathbf{x}) \;=\; (2\pi\hat{\sigma}^2)^{-n/2} e^{-\frac{n}{2}}$$

So

$$\lambda(\mathbf{x}) \;=\; \left(\frac{\hat{\sigma}^2}{\sigma_0^2}\right)^{\frac{n}{2}} e^{-\frac{1}{2\sigma_0^2}\sum_{i=1}^{n}(x_i-\bar{x})^2+\frac{n}{2}}$$

Put

$$t = \frac{1/n\sum_{i=1}^{n}(x_i-\mu)^2}{\sigma_0^2}$$

then

$$\lambda(\mathbf{x}) = t^{n/2}\exp\left\{\frac{n}{2}t + \frac{n}{2}\right\}$$

and

$$-2\log\lambda = n(t - 1 - \log t)$$

increases as $t$ increases from 1, decreases as $t$ decreases from 1, so:

reject $H_0$ if $|t-1|$ large

Under $H_0$, $T = t(\mathbf{X}) = \frac{(n-1)S^2}{\sigma_0^2} \sim \chi_n^2$

so two-tailed test at level $\alpha$: reject $H_0$ is $t >$ $\chi_n^2(\alpha/2)$ or $t < \chi_n^2(1-\alpha/2)$

*Example*: Multinomial distribution:

$n$ balls, $m$ cells, probability $p_i$ for cell $i$, $i = 1, \ldots, m$, balls thrown independently

$x_i$ is the count in cell $i$, $i = 1, \ldots, m$

$$f(\mathbf{x}, (p_1, \ldots, p_m; m)) = \binom{n}{x_1, \ldots, x_m} p_1^{x_1} \cdots p_m^{x_m}$$

*Exercise:* M.l.e. if $\widehat{p}_i = \frac{x_i}{n}$, $i = 1, \ldots, m$

$H_0 : p_i = p_i(\theta)$, $i = 1, \ldots, m$

$H_1 : not$

$$\begin{aligned} \lambda(\mathbf{x}) &= \frac{p_1(\widehat{\theta})^{x_1} \cdots p_m(\widehat{\theta})^{x_m}}{\widehat{p}_1^{x_1} \cdots \widehat{p}_m^{x_m}} \\ &= \prod_{i=1}^{m} \left( \frac{p_i(\widehat{\theta})}{\widehat{p}_i} \right)^{x_i} \end{aligned}$$

and

$$
\begin{aligned}
-2 \log \lambda(\mathbf{x}) &= -2n \sum_{i=1}^{m} \widehat{p}_i \log \left( \frac{p_i(\widehat{\theta})}{\widehat{p}_i} \right) \\
&= 2 \sum_{i=1}^{m} O_i \log \left( \frac{O_i}{E_i} \right),
\end{aligned}
$$

where $O_i = n\widehat{p}_i = x_i$ observed count in cell $i$

$E_i = np_i(\widehat{\theta})$ expected count in cell $i$ if $H_0$ is true

If $\theta \in \mathbf{R}^k$ then $dim\Theta = m - 1, dim\Theta_0 = k$, so under $H_0$

in distribution

$$-2 \log \lambda(\mathbf{x}) \approx \chi^2_{m-1-k}$$

Taylor expansion:

$$x \log\left(\frac{x}{x_0}\right) \approx x - x_0 + \frac{1}{2}(x - x_0)^2 \frac{1}{x_0}$$

so

$$\begin{aligned} & -2\log\lambda(\mathbf{x}) \\ \approx\ & 2n\sum_{i=1}^{m}(\widehat{p}_i - p_i(\widehat{\theta})) + n\sum_{i=1}^{m}\frac{(\widehat{p}_i - p_i(\widehat{\theta}))^2}{p_i(\widehat{\theta})} \\ =\ & \sum_{i=1}^{m}\frac{(x_i - np_i(\widehat{\theta}))^2}{np_i(\widehat{\theta})} \\ =\ & \sum_{i=1}^{m}\frac{(O_i - E_i)^2}{E_i} \end{aligned}$$

is called *Pearson's Chisquare-Statistic*

*Examples*:

six-faced die, $m = 6$, $H_0 : p_i = 1/6, i = 1, \ldots, m$

"the die is fair"

$H_0 : p_i = e^{-\lambda} \frac{\lambda^i}{i!}, i = 1, \ldots, m - 1$ "Poisson"

Can be used to assess goodness of fit of distributional assumptions: next chapter.