STAT 249
Lecture 1.7
3/4/16

Confidence Intervals

recap:

Confidence Interval:

a random interval $[\hat{L}, \hat{U}]$ that includes the "state of nature" θ with some probability $1-\alpha$ (in the example last time, 95%).

recap: **Testing Composite Hypotheses**

H_0 : Group of θ_0's
H_1 : Group of θ_1's
} mutually exclusive

Likelihood Ratio Test

Reject H_0 if $\lambda < \lambda_c$ $(0 \leq \lambda \leq 1)$

$$\lambda = \frac{\max\limits_{\theta_0\text{'s}} L(\theta)}{\max\limits_{\text{all } \theta\text{'s}} L(\theta)}$$

$\alpha = P(\text{Reject } H_0 / H_0 \text{ true})$ "Type 1 error" "False negative"

$\beta = P(\text{Accept } H_0 / H_1 \text{ true})$ "Type 2 error" "False positive"

$\pi = 1 - \beta$ = power of the test

Ex

Consider $X_1, \dots, X_n$ iid observations from a normal distribution having unknown mean μ and known variance σ^2.

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

Suppose we have a test at false positive level α that rejects if

$$|\bar{X} - \mu_0| > x_0$$

where we picked x_0 so that

if H_0 is true, $P(|\bar{X} - \mu_0| > x_0) = \alpha$

What is x_0? denote the standard deviation of $\bar{X}$ by $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$.

This is a two sided test, so we want $x_0 = \sigma_{\bar{x}}\, z\left(\frac{\alpha}{2}\right)$, $z = \frac{\bar{x} - \mu}{\sigma}$

so the test will accept when

$$|\bar{X} - \mu_0| < \sigma_{\bar{x}} \left(\frac{\alpha}{2}\right)$$

Getting rid of the absolute values by writing positive and negative cases explicitly, we now have

$$-\sigma_{\bar{X}} z\left(\frac{\alpha}{2}\right) < \bar{X} - \mu_0 < \sigma_{\bar{X}} z\left(\frac{\alpha}{2}\right)$$

$$\bar{X} - \sigma_{\bar{X}} z\left(\frac{\alpha}{2}\right) < \mu_0 < \bar{X} + \sigma_{\bar{X}} z\left(\frac{\alpha}{2}\right)$$

This is almost exactly the same expression we got at the end of lec 16. That was a confidence interval for the <u>estimation</u> of μ. Here we see that <u>if</u> the hypothesis test accepts H_0 at level α, the $100(1-\alpha)\%$ confidence interval for μ_0 is

$$\left[\bar{X} - \sigma_{\bar{X}} z\left(\frac{\alpha}{2}\right), \bar{X} + \sigma_{\bar{X}} \left(\frac{\alpha}{2}\right)\right]$$

So the confidence interval is precisely those values of μ_0 for which $H_0: \mu = \mu_0$ is accepted.

This is true in general:

Thm Suppose for every value θ_0 in Θ there is a test at level α of the hypothesis $H_0: \theta = \theta_0$. Denote the acceptance region of the test by $A(\theta_0)$. Then the set

$$C(\vec{X}) = \{\theta : \vec{X} \text{ in } A(\theta)\}$$

is a $100(1-\alpha)\%$ confidence region for θ.

Pf $P[\vec{X} \text{ in } A(\theta_0) \mid \theta = \theta_0] = 1-\alpha$

by definition of $A(\theta)$.

$$P[\theta_0 \text{ in } C(\vec{X}) \mid \theta = \theta_0] = P[X \text{ in } A(\theta_0) \mid \theta = \theta_0]$$

by definition of $C(\vec{X})$.

So, a $100(1-\alpha)\%$ confidence region for θ consists of all those values of θ_0 for which $H_0: \theta = \theta_0$ will not be rejected at level α.

True the other way also;

Thm

Suppose $C(\vec{X})$ is a $100(1-\alpha)\%$ confidence region for θ; in other words for every θ_0,

$$P[\theta_0 \in C(\vec{X}) \mid \theta = \theta_0] = 1 - \alpha$$

Then the acceptance region for a test at level α of the hypothesis $H_0 : \theta = \theta_0$ is

$$A(\theta_0) = \{\vec{X} \mid \theta_0 \in C(\vec{X})\}$$

Pf The test has level α because

$$P(\vec{X} \in A(\theta_0) \mid \theta = \theta_0) = P(\theta_0 \text{ in } C(\vec{x}) \mid \theta = \theta_0)$$
$$= 1 - \alpha$$

That is,

$H_0 : \theta = \theta_0$ is accepted if θ_0 lies in the confidence region.

Multinomial and Poisson

Consider a set of indep random vars $(x_1, \ldots, x_n)$ $\quad X \sim \text{Poisson}(\lambda)$

$H_0 : \lambda_i = \lambda$ for all i (rates the same)

$H_1 : \lambda_i \neq \lambda_k \quad i \neq k$ (rates different)

For an indep Poisson r.v.,

$$f(x|\theta) = f(x|\lambda) = \frac{\lambda^x}{x!} e^{-\lambda}$$

note; we are denoting M.C. Poisson "state of nature" by λ, not θ

Denote the MLE for H_0 by $\hat{\lambda} = \overline{X}$

Denote the MLEs for H_1 by $\tilde{\lambda}_1 = \hat{\lambda}_1, \ldots, \tilde{\lambda}_n$. $\tilde{\lambda}_1 = x_1, \ldots, \tilde{\lambda}_n = x_n$

The liklihood ratio is then (write Λ to avoid confusion):

$$\Lambda = \frac{\max_{\lambda} L(\lambda)}{\max_{\text{all } \lambda_i\text{'s}} L(\lambda)} = \frac{\prod_{i=1}^{n} \hat{\lambda}^{x_i} \frac{e^{-\hat{\lambda}}}{x_i!}}{\prod_{i=1}^{n} \tilde{\lambda}_i^{x_i} \frac{e^{-\tilde{\lambda}_i}}{x_i!}}$$

$$= \prod_{i=1}^{n} \left(\frac{\hat{\lambda}}{\tilde{\lambda}_i}\right)^{x_i} e^{\lambda_i - \lambda} = \prod_{i=1}^{n} \left(\frac{\bar{x}}{x_i}\right)^{x_i} e^{x_i - \bar{x}}$$

We had the LR

$$\Lambda = \prod_{i=1}^{n} \left(\frac{\bar{x}}{x_i}\right)^{x_i} e^{x_i - \bar{x}}$$

$$\log \Lambda = \sum_{i=1}^{n} \left(x_i \log\left(\frac{\bar{x}}{x_i}\right) + (x_i - \bar{x})\right)$$

$$= -\sum_{i=1}^{n} x_i \log\left(\frac{x_i}{\bar{x}}\right)$$

But wait! We've seen this before, Lec 14, pp 14-15. The L.R. for the multinomial dist was (write LR as Λ)

$$\log \Lambda = \sum_{i=1}^{n} x_i \log\left(\frac{x_i}{m_i}\right)$$

x_i observed
m_i expected from multinomial dist

there, we found that

$$-2 \log \Lambda \stackrel{\sim}{=} \sum_{i=1}^{n} \frac{(x_i - m_i)^2}{m_i} = \chi^2$$

with Poisson,

$$-2 \log \Lambda \stackrel{\sim}{=} \frac{1}{\bar{x}} \sum_{i=1}^{n} (x_i - \bar{x})^2 = \chi^2$$

under H_0, there is one parameter $\hat{\lambda}$, so d.f. $= n-1$, just like multinomial.

Now, the estimated variance

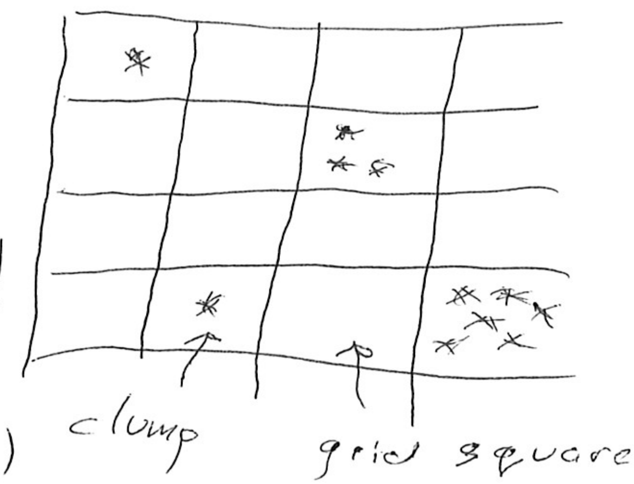
$$s^2 = \frac{1}{(n-1)} \sum (x_i - \bar{x})^2,$$

so

$$-2 \log \Lambda \approx \frac{1}{\bar{x}} \sum_{i=1}^{n} (x_i - \bar{x})^2 \propto \frac{\text{var}}{\text{mean}}$$

For Poisson r.v.'s, $\sigma^2 = \mu$, so deviations from this ratio are being tested.

<u>Ex</u> clumps of bacteria
0.01 ml milk spread out on slide with grid:



There aren't very many bacteria in the milk (fortunately!), and one often is told that Poisson statistics are useful for "rare events" (e.g. death by Horsekick).

So let us consider some actual data (Bliss and Fisher, Biometrics <u>9</u>: 174-200)

The clump data: (from 400 squares)

# bacteria/sq.	0	1	2	3	4	5	6	7	8	9	10	19
Frequency	56	104	80	62	42	27	9	9	5	3	2	1

$$\hat{\lambda} = \bar{x} = \frac{0(56) + 1(104) + 2(80) + \cdots + 19(1)}{400}$$

$$= 2.44$$

Observed	56	104	80	62	42	27	9	20
Expected	34.9	85.1	103.8	84.4	51.5	25.1	10.2	5.0
χ^2 contrib	12.8	4.2	5.5	5.9	1.8	.14	.14	45.0

$\chi^2_6 = 75.4$, $\quad$ d.f. $= \underset{\text{cells}}{8} - \underset{\text{for Poisson or multinomial}}{1} - \underset{\text{parameter estimated}}{1} = 6$

6 d.f

p-value << .005 ... reject!

In general, χ^2 is a one-sided test - reject H_0 if $\chi^2 > c$

But small χ^2 is sometimes informative, too.

Mendel's Peas (test for indep, multinomial dist)

We expect:

	Smooth	Wrinkled
yellow	$\frac{9}{16}$	$\frac{3}{16}$
green	$\frac{3}{16}$	$\frac{1}{16}$

For a particular experiment, Mendel reported (compare to observed, calculated from above table)

Type	Observed	Expected
smooth yellow	315	312.75
smooth green	108	104.25
wrinkled yellow	102	104.25
wrinkled green	31	34.75

mmm... looks like good agreement!

d.f. = 3. $\chi^2 = 0.604$ (G test gives 0.618)

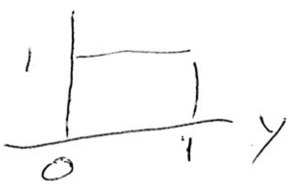
p-value close to 0.9. Can't reject. But would expect agreement <u>worse</u> than this 90% of the time... means one should look at Mendel's other experiments.

To do this, we need a way of combining results of many experiments ("meta analysis"), and we'll need to (somehow) combine their p-values.

We'll need (also see Lec. 3, p. 15)

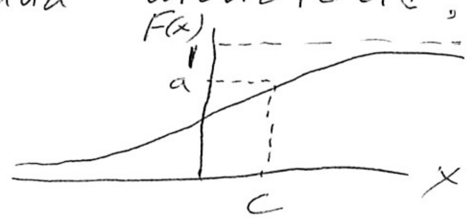
<u>Thm</u> X a continuous rand. var. with c.d.f. $F(x)$.

$Y = F(X)$ (transform of X).



<u>Then</u> $Y \sim \text{Uniform}[0,1]$

<u>Pf</u> $F(x)$ continuous and monotone:



$$P(Y \le y) = P(F(X) \le y)$$
$$= P(X \le F^{-1}(y))$$
$$= F(F^{-1}(y)) = y$$

$a = F(c)$, $c = F^{-1}(a)$

hence density of Y is $= \begin{cases} 1 & 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$

Corollary: Test H_0 with test statistic T, where T has a cont. dist., with c.d.f. F_0 under H_0, and we reject for small T. Then if $T = t$ is observed, $F_0(t) = P$-value, smallest α for which you reject H_0

Then $P = F_0(T)$ is the P-value and has a uniform $(0,1)$ dist under H_0.

[Family of tests: Reject if $T \le c$, small c, small α. Smallest $= T$]

Combining Indep Tests

k indep tests of some H_0

P-values $P_1, P_2, \ldots, P_k$

Do these, "combined", provide evidence to reject H_0? What is the "combined" P-value?

<u>One idea</u>: $P = P_1 P_2 P_3 \ldots P_k$ (product)

<u>not</u> right - way too small ($\to 0$ as $k \to \infty$) and indep tests of the <u>same</u> hypothesis not completely indep.

But look at $\underline{-2 \log P}$. Claim:

This has a χ^2_{2k} dist under H_0 (if dists continuous).

Remember that χ^2_k has pdf

$$f(x) = \frac{1}{2^{\frac{k}{2}}\Gamma(\frac{k}{2})} x^{\frac{k}{2}-1} e^{-\frac{x}{2}}$$

$$-2 \log P = \sum_1^n -2 \log P_i.$$ Let $Y_i = -2 \log P_i$

Under H_0, $$P(Y \le y) = P(-2 \log P_i \le y)$$
$$= P(P_i \ge e^{-y/2})$$
$$= 1 - e^{-y/2}$$

So under H_0, Y is $\underline{\text{exponential}}$ density $\frac{1}{2} e^{-y/2}$

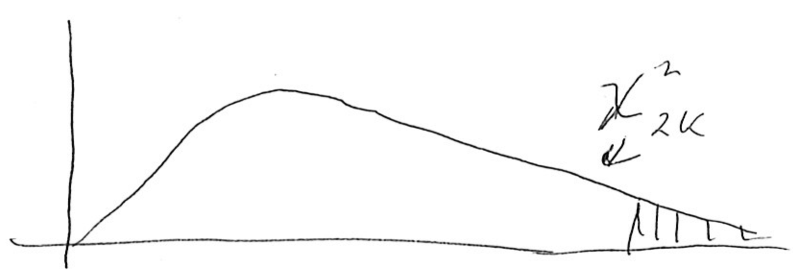
→ see above χ^2_2 is exponential !!

So

$$-2 \log P = \chi^2_2 + \chi^2_2 + \cdots + \chi^2_2$$
$$= (U_1^2 + U_2^2) + (U_3^2 + U_4^2) + \cdots$$

dist χ^2_{2k}

P small, $-2 \log P$ large



Examples:

① P_i's: 0.3, 0.4, .5, .6, .7

$$P = \prod_{i=1}^{5} P_i = 0.025$$

$$-2 \log P = 7.36$$

$k = 5$

$2k = 10$ d.f.

(P-value = .69)

② P_i's: 0.1, 0.15, 0.08, 0.2, 0.07

$$P = .000017$$

$$-2 \log P = 21.99$$

"Fisher's Method of Combination"

(metaanalysis)

Done on Mendel's experiments, Fisher got $\underbrace{\text{P-value of 0.99996.}}$ (!?!)

Mendel's laws are correct, but Mendel's variance is awfully small. Probably he kept the best results, not knowing principles of statistics, not yet invented when he did his work.