Στατιστική και πιθανότητες

Μέρος ΙΙ

Στατιστικοί εκτιμητές

- Όταν κάνουμε στατιστική ανάλυση, ένας στόχος είναι η αληθής εκτίμηση της τιμής μίας ή περισσότερων παραμέτρων από πειραματικά δεδομένα και η κατανόηση της αβεβαιότητας της μέτρησης αυτής

- Σημαντικά χαρακτηριστικά ενός καλού εκτιμητή είναι:
  - Συνέπεια: Αν ο όγκος των δεδομένων είναι μεγάλος, η εκτίμηση συγκλίνει στην πραγματική τιμή:
    - Bias: Αν υπάρχει διαφορά μεταξύ της αναμενόμενης τιμής της εκτίμησης της παραμέτρου και της πραγματικής τιμής της παραμέτρου
  - Ανθεκτικότητα (robustness): Η εκτίμηση δεν αλλάζει ιδαίτερα αν η πραγματική pdf (probability density function) διαφέρει από την υποτιθέμενη pdf. Για παράδειγμα στις ουρές μίας κατανομής.

- Χρειαζόμαστε ακόμα να ξέρουμε την αβεβαιότητα της εκτίμησής μας, δηλαδή πόσο μακριά είναι η εκτιμούμενη από την πραγματική τιμή εξαιτίας στατιστικών διακυμάνσεων στο σύνολο των μετρήσεων

Η μέθοδος των ελαχίστων τετραγώνων

- Υποθέτουμε ότι οι μετρήσεις μας προέρχονται από μεγάλη στατιστική και επομένως μπορούμε να υποθέσουμε ότι βρισκόμαστε στη Gaussian περιοχή
- Θέλουμε να έχουμε τις καλύτερες εκτιμήσεις των παραμέτρων της συνάρτησης που περιγράφει τα δεδομένα
- Το επιτυγχάνουμε ελαττώνοντας τη διασπορά των δεδομένων από τη συνάρτηση προσαρμογής λαμβάνοντας υπόψην την αβεβαιότητα των δεδομένων:
- Η διασπορά προσδιορίζεται συναρτήσει μίας ποσότητας: $\chi^2 = \sum_{i=1}^{N} \frac{\left(x_i - \mu\right)^2}{\sigma^2}$
- Μπορούμε να γράψουμε το χ^2 συναρτήσει των παρατηρούμενων μεγεθών:

$$\chi^2 = \sum_{i=1}^{N} \frac{\left(y_i - F\left(x_i, \theta\right)\right)^2}{\sigma_i^2}$$

- Ελαχιστοποιούμε το χ^2 ως προς την/τις παραμέτρους θ_i
- Ιδιαίτερα χρήσιμη στις περιπτώσεις δειγμάτων μεγάλης στατιστικής όπου η σύγκλιση της πολλών της αρνητικής likelihood ($-\ln\mathcal{L}$) είναι αργή

Συνάρτηση μέγιστης πιθανότητας – Likelihood

- Η likelihood, $\mathcal{L}(x;\theta)$ εκφράζει την πιθανότητα ότι μία μέτρηση του *x,* θα δώσει συγκεκριμένη τιμή για κάποια δεδομένη θεωρία.
  - Για να προσδιορίσουμε την likelihood, θα πρέπει να γνωρίζουμε τόσο τη θεωρία όσο και τις τιμές των παραμέτρων από τις οποίες εξαρτάται η θεωρία
- Αν έχουμε ένα σύνολο μετρήσεων, η ολική likelihood βρίσκεται από το γινόμενο των likelihoods των μετρήσεων:

$$\mathcal{L}(x;\theta) = \prod_{i=1}^{N} \mathcal{L}_i(x;\theta)$$

όπου *θ* μπορεί να αναπαραστά μία ή περισσότερες παραμέτρους

Φυσικός λογάριθμος πιθανότητας – Log Likelihood

- Για να εκτιμήσουμε την/τις παραμέτρους *θ*, θα πρέπει να μεγιστοποιήσουμε την πιθανότητα
- Συνηθισμένη τεχνική για να βρούμε το μέγιστο είναι να θέσουμε την παράγωγο της ποσότητας που μεγιστοποιούμε στο 0
- Είναι ευκολότερο να μεγιστοποιήσουμε το λογάριθμο της πιθανότητας $\ln \mathcal{L}$

$$\frac{\partial \ln \mathcal{L}}{\partial \theta} = \frac{\partial}{\partial \theta} \ln \prod_{i=1}^{N} \mathcal{L}_i = \frac{\partial}{\partial \theta} \sum_{i=1}^{N} \ln \mathcal{L}_i = 0$$

- Το μέγιστο της likelihood συνάρτησης ατιστοιχεί σε τιμές της/των παραμέτρων *θ*, που ελαχιστοποιούν την ποσότητα: $-\sum_{i=1}^{N} \ln \mathcal{L}_i$
- Αν υπάρχουν πολλές παράμετροι *θ*, μπορούμε να ελαχιστοποιήσουμε ως προς κάθε μία από αυτές
  - Θα εξετάσουμε πιθανές συσχετίσεις μεταξύ τους

Παράδειγμα Poisson κατανομής για likelihood

- Έστω N ανεξάρτητες προσπάθειες, η κάθε μία με αποτέλεσμα n_i
- Η likelihood συνάρτηση να παρατηρήσουμε n_i όταν η παραγματική τιμή είναι μ

$$\mathcal{L}_i\left(n_i;\mu\right)=\frac{e^{-\mu}\left(\mu\right)^{n_i}}{n_i!}$$

- Εφόσον έχουμε N μετρήσεις, το γινόμενο των likelihoods για τις N μετρήσεις είναι:

$$\mathcal{L}\left(data;\mu\right)=\prod_{i=1}^{N}\frac{e^{-\mu}\left(\mu\right)^{n_i}}{n_i!}\Rightarrow\ln\mathcal{L}=\sum_{i=1}^{N}\ln\left(\frac{e^{-\mu}\left(\mu\right)^{n_i}}{n_i!}\right)$$

$$\Rightarrow\ln\mathcal{L}=\sum_{i=1}^{N}\left(-\mu+n_i\ln\left(\mu\right)-\ln\left(n_i!\right)\right)\Rightarrow\ln\mathcal{L}=-N\mu+\ln\left(\mu\right)\sum_{i=1}^{N}\left(n_i\right)+const$$

- Παραγωγίζουμε την τελευταία ως προς τις παραμέτρους της θεωρίας, $\theta = \mu$

$$\left.\frac{d\ln\mathcal{L}}{d\mu}\right|_{\hat{\mu}=\mu}=-N+\frac{\sum_{i=1}^{N}n_i}{\mu}=0\ \Rightarrow\hat{\mu}=\frac{1}{N}\sum_{i=1}^{N}n_i$$

- Όπως αναμένονταν, η καλύτερη εκτίμηση για την πραγματική τιμή είναι η μέση τιμή

Likelihood Gaussian κατανομής

- Έστω η Gaussian κατανομή: $G\left(x;\mu,\sigma\right)=\frac{1}{\sqrt{2\pi\sigma}}e^{-\frac{\left(x-\mu\right)^2}{2\sigma^2}}$

- Παίρνουμε την παράγωγο της log likelihood συνάρτησης:

$$\frac{\partial}{\partial\mu}\left(\ln\mathcal{L}\right)\Big|_{\hat{\mu}=\mu}=\frac{\partial}{\partial\mu}\left(-\sum_{i=1}^{N}\frac{\left(x_i-\mu\right)^2}{2\sigma^2}+const\right)\Rightarrow\frac{\partial}{\partial\mu}\left(\ln\mathcal{L}\right)\Big|_{\hat{\mu}=\mu}=-\sum_{i=1}^{N}\frac{\left(x_i-\mu\right)}{\sigma^2}\Big|_{\hat{\mu}=\mu}=0$$

$$\Rightarrow\hat{\mu}=\frac{1}{N}\sum_{i=1}^{N}x_i$$

- Η εκτίμηση για την αβεβαιότητα σ, είναι: $\sigma=\frac{1}{N-1}\sum_{i=1}^{N}\left(x_i-\mu\right)^2$

Binned ή unbinned log likelihood

- Ο φορμαλισμός likelihood δουλεύει για οποιαδήποτε καλά συμπεριφερόμενη pdf
- Το γινόμενο της likelihood είναι ένα γινόμενο ως προς τις μετρήσεις
- Μπορούμε να προσδιορίσουμε την έννοια της μέτρησης
- Παράδειγμα: Η μέτρηση του χρόνου ζωής ενός σωματιδίου από ένα σύνολο τέτοιων σωματιδίων που παράγονται τη χρονική στιγμή t=0 και διασπώνται σε χρόνο t:

$$f(t) = \frac{1}{\tau} e^{-t/\tau}$$

Δύο τρόποι για να κατασκευάσουμε τη συνάρτηση likelihood:

1) Για κάθε διάσπαση i μετρούμε το χρόνο t_i. Κατασκευή της likelihood από το γινόμενο όλων των μετρούμενων χρόνων (unbinned likelihood)
2) Κατασκευή ενός histogram του αριθμού των διασπάσεων σε bins του χρόνου. Στην περίπτωση αυτή, η μέτρηση είναι ο αριθμός των διασπάσεων σε κάθε bin i (binned likelihood)

Σύνδεση log likelihood και χ^2

- Είδαμε προηγουμένως ότι για την περίπτωση της Gaussian κατανομής:

$$\ln \mathcal{L} = -\sum_{i=1}^{N} \frac{\left(x_i - \mu\right)^2}{2\sigma^2} + const$$

- Η σχέση αυτή συγκρίνεται με την αντίστοιχη του χ^2: $\chi^2 \equiv \sum_{i=1}^{N} \frac{\left(x_i - \mu\right)^2}{\sigma^2}$

- Αντιστοιχία μεταξύ των δύο σχέσεων προκύπτει ότι: $\chi^2 \equiv -2 \ln \mathcal{L}$

- Ο φορμαλισμός της likelihood δουλεύει για όλες τις pdf και επομένως είναι περισσότερος γενικός και θα πρέπει να χρησιμοποιείται γενικά

Log likelihood

- Ο λογάριθμος της likelihood συνάρτησης μπορεί να αναπτυχθεί κατά Taylor ως προς το ελάχιστό της:

$$\left.\frac{\partial \ln \mathcal{L}}{\partial \theta}\right|_{\theta=\theta_{\min}} = 0 \Rightarrow \ln \mathcal{L} = \ln \mathcal{L}_{\min} + \frac{1}{2}\left.\frac{\partial^2 \ln \mathcal{L}}{\partial \theta^2}\right|_{\theta=\theta_{\min}} \left(\theta - \theta_{\min}\right)^2$$

$$\Rightarrow 2\left(\ln \mathcal{L} - \ln \mathcal{L}_{\min}\right) = \left.\frac{\partial^2 \ln \mathcal{L}}{\partial \theta^2}\right|_{\theta=\theta_{\min}} \left(\theta - \theta_{\min}\right)^2$$

- Στο όριο του μεγάλου αριθμοπθυ μετρήσεων *N,* η κατανομή της $\mathcal{L}$ χάρη στο κεντρικό θεώρημα (θεώρημα μεγάλων αριθμών) γίνεται Gaussian.
- Για Gaussian κατανομές αλλαγή στην $2\ln\mathcal{L}$ κατά μία μονάδα, ισοδυναμεί σε αλλαγή κατά 1σ στην παράμετρο *θ.* Επομένως η αβεβαιότητα στη *θ* είναι:

$$\sigma_\theta^2 \equiv \left\langle \left(\theta - \theta_{\min}\right)^2 \right\rangle = -\frac{1}{\dfrac{\partial^2 \ln \mathcal{L}}{\partial \theta^2}}$$

- Η αβεβαιότητα στις εκτιμούμενες τιμές των παραμέτρων *θ* μπορούν να βρεθούν υπολογίζοντας την τιμή του Δ*θ* για την οποία το $-2\ln\mathcal{L}$ αλλάζει κατά μία μονάδα . Αν η $\ln\mathcal{L}$ δεν είναι παραβολική, η αβεβαιότητα είναι ασυμμετρική

Παράδειγμα

- Έστω ότι κάναμε μία σειρά από μετρήσεις x_i. Έστω ακόμα ότι ο αριθμός των γεγονότων συναρτήσει του x ακολουθεί την κατανομή:

$$N(x) = A + Bx \qquad \text{για} \quad 0 < x < 10$$

Θα χρησιμοποιήσουμε τη μέθοδο της likelihood για να εκτιμήσουμε το λόγο $k = A/B$

- Πώς θα πρέπει να χτίσουμε το πρόβλημα:
  - Ο συνολικός αριθμός των γεγονότων N_{tot} μπορεί να υπολογισθεί από

$$N_{tot} = \int_0^{10} (A + Bx)\,dx \Rightarrow N_{tot} = \left(Ax + \frac{1}{2} Bx^2 \right)\Bigg|_0^{10} \Rightarrow N_{tot} = 10A + 50B$$

  - Κανονικοποίηση της συνάρτησης πυκνότητας πιθανότητας:

$$f(x; A, B) = \frac{1}{N_{tot}}(A + Bx) = \frac{1}{10A + 50B}(A + Bx) = \frac{A}{10A + 50B} + \frac{Bx}{10A + 50B}$$

$$f(x; A, B) = 0.1\left(\frac{k}{k+5} + \frac{x}{k+5} \right)$$

  - Η likelihood συνάρτηση θα είναι: $\mathcal{L}(x; k) = \prod_{i=1}^{N} 0.1\left(\frac{k}{k+5} + \frac{x}{k+5} \right)$

Παράδειγμα

- Η log likelihood συνάρτηση θα είναι επομένως:

$$\ln\left(\mathcal{L}\left(x;k\right)\right)=\sum_{i=1}^{N}\ln\left[0.1\left(\frac{k}{k+5}+\frac{x}{k+5}\right)\right]$$

$$\Rightarrow \ln\left(\mathcal{L}\left(x;k\right)\right)=\sum_{i=1}^{N}\ln\left[\left(\frac{k}{k+5}+\frac{x}{k+5}\right)\right]+N\ln\left(0.1\right)$$

- Αν οι τιμές των x_i είναι γνωστές τότε μπορούμε να ελαχιστοποιήσουμε την $-\ln\mathcal{L}$ ως προς τον λόγο *k*. Ο τελευταίος όρος είναι σταθερός και ανεξάρτητος του *x* απλά προσθέτει ένα σταθερό όρο στη log likelihood και δεν παίζει ρόλο στην ελαχιστοποίηση.
- Η ελαχιστοποίηση μπορεί να γίνει με διάφορα λογισμικά (όπως το ROOT που περιέχει το λογισμικό ελαχιστοποίησης Minuit) .
- Ωστόσο το ελάχιστο μπορεί να βρεθεί παρατηρώντας πώς αλλάζει η $\ln\mathcal{L}\left(x;k\right)$ καθώς μεταβάλεται η παράμετρος *k*.
- Μπορείτε να κατεβάσετε το πρόγραμμα από το εργαστήριο 10 για το ROOT το οποίο δημιουργεί δεδομένα για A=1 και B=2 και τα χρησιμοποιεί για να βρεί το *k*.

Σχετιζόμενες μεταβλητές

- Σε πολλές περιπτώσεις, οι μεταβλητές που χρησιμοποιούνται σε προσαρμογή δεδομένων δεν είναι ανεξάρτητες μεταξύ τους
- Κατά τη διάρκεια της ελαχιστοποίησης θα πρέπει να λαμβάνονται υπόψην συσχετίσεις μεταξύ των παραμέτρων
- Σαν υπενθύμιση, η απόκλιση των μετρήσεων δίνεται από τη σχέση:

$$\sigma^2 \equiv Var(x) = \int_{-\infty}^{+\infty} x^2 f(x)\,dx - \mu^2$$

- Ορίζουμε την covariance σχέση, Cov[x,y], μεταξύ δύο μεταβλητών *x* και *y*:

$$\mathrm{cov}[x,y] = \int_{-\infty}^{+\infty} xy f(x,y)\,dx\,dy - \mu_x \mu_y$$

- Αν οι δύο μεταβλητές *x* και *y* είναι ανεξάρτητες μεταξύ τους, τότε:

$$\mathrm{cov}[x,y] = 0 \quad \text{για } x \neq y \text{ (ανεξάρτητα)}$$

Περισσότερες των δύο σχετιζόμενες μεταβλητές

- Κάθε μέτρηση είναι ένα σύνολο N ποσοτήτων, x_1, x_2, … , x_N
- Ο πίνακας συσχέτισης, covariance matrix, είναι ένας N x N πίνακας με στοιχεία:

$$V_{ij} = \mathrm{cov}\left[x_i, x_j\right] \equiv \left\langle \left\langle x_i - \mu_i \right\rangle \left\langle x_j - \mu_j \right\rangle \right\rangle$$

- Για μή σχετιζόμενες (ανεξάρτητες) μεταβλητές, ο πίνακας αυτός είναι διαγώνιος
- Μία σχετική ποσότητα που χρησιμοποιούμε αρκετές φορές είναι ο παράγοντας συσχέτισης:

$$\rho_{ij} = \frac{V_{ij}}{\sqrt{V_{ii} V_{jj}}} \equiv \frac{V_{ij}}{\sigma_i \sigma_j}$$

- Οι τιμές του παράγοντα συσχέτισης βρίσκονται στο διάστημα: $-1 \le \rho_{ij} \le 1$

Ο covariance πίνακας για την Gaussian περίπτωση

- Έστω *x* και *y* ανεξάρτητες μεταβλητές

$$G\left(x,y;\mu_x,\sigma_x,\mu_y,\sigma_y\right)=\frac{1}{\sqrt{2\pi\sigma_x}}e^{-\frac{\left(x-\mu_x\right)^2}{2\sigma_x^2}}\frac{1}{\sqrt{2\pi\sigma_y}}e^{-\frac{\left(y-\mu_y\right)^2}{2\sigma_y^2}}$$

  - Η 2η παράγωγος ως προς μ δίνει: $\frac{\partial^2}{d\mu_x^2}\left(\ln\mathcal{L}\right)=-\sum_{i=1}\frac{1}{\sigma_x^2}$

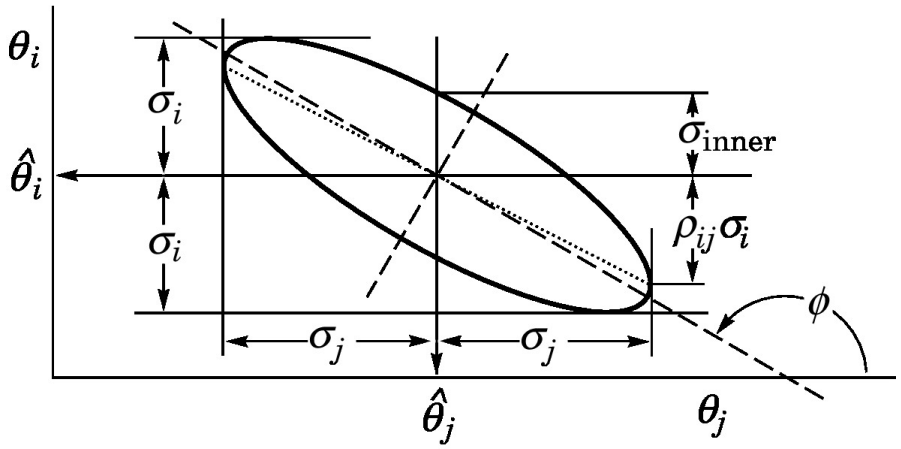
- Υποθέτουμε τώρα ότι τα *x* και *y* δεν είναι ανεξάρτητες μεταξύ τους
- Αντιστρέφουμε τον covariance πίνακα για μια ομάδα εκτιμητών μέγιστης πιθανότητας που ορίζονται σύμφωνα με τη:

$$\left\langle \hat{V}^{-1}\right\rangle_{ij}=-\frac{\partial^2\ln\mathcal{L}}{\partial\mu_i\,\partial\mu_j}$$

- Για την περίπτωση της binned likelihood, και σε περιοχή με μεγάλο αριθμό N μετρήσεων, η likelihood μπορεί να προσεγγισθεί με χ^{2}¨

$$\left\langle \hat{V}^{-1}\right\rangle_{ij}=\frac{1}{2}\frac{\partial^2\chi^2}{\partial\mu_i\,\partial\mu_j}$$

Αποτέλεσμα συσχετιζόμενων αβεβαιοτήτων



- Έλλειψη σφάλματος για δύο παραμέτρους οι οποίες έχουν αρνητική συσχέτιση
- Η κλίση σχετίζεται με τον παράγοντα συσχέτισης: $d\theta_i \big/ d\theta_j$ όπου θ είναι εκτιμητές (στο προηγούμενο παράδειγμα ήταν οι μέσες τιμές)
- Ο correlation πίνακας συνήθως προσδιορίζεται από τα δεδομένα κατά τη διάρκεια της διαδικασίας προσαρμογής.

Διάδοση σφαλμάτων

- Κάποιος μπορεί να ανατρέξει στο internet:
  http://en.wikipedia.org/wiki/Propagation_of_uncertainty

- Η βασική σχέση που χρησιμοποιούμε είναι:

$$\sigma_f^2 = \left(\frac{\partial f}{\partial a}\right)^2 + \left(\frac{\partial f}{\partial \beta}\right)^2 + 2\left(\frac{\partial f}{\partial a}\right)\left(\frac{\partial f}{\partial \beta}\right)\mathrm{cov}\left[a,\beta\right]$$

  όπου στο συγκεκριμένο μοντέλο υπάρχουν δύο παράμετροι α και β

- Η σχέση μπορεί να επεκταθεί σε περισσότερες διαστάσεις και συνήθως εκφράζεται με τη βοήθεια ενός πίνακα

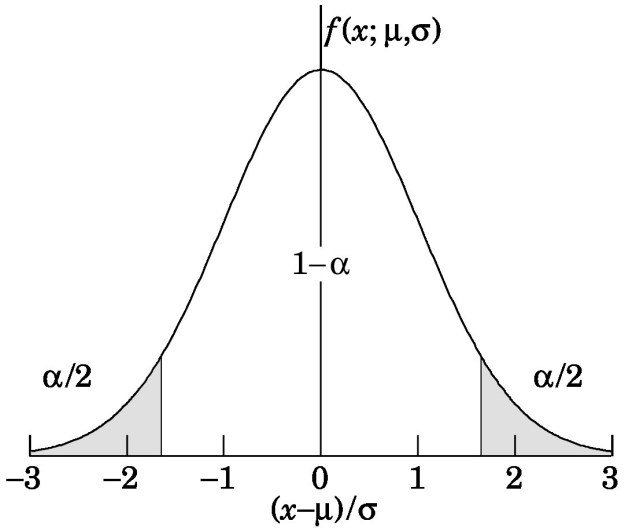
- Αν οι παράμετροι είναι μή σχετιζόμενες (ανεξάρτητες μεταξύ τους) τότε η σχέση περιορίζεται σε αυτή που έχετε δει στα εισαγωγικά εργαστήρια

Διαστήματα εμπιστοσύνης – confidence intervals

- Χρησιμοποιώντας τη γλώσσα των frequists: το ποσοστό του αποτελέσματος το οποίο δεν περιέχεται μεταξύ x_L και x_U είναι:

$$1-a=\int_{x_L}^{x_U} P\left(x;\theta\right)dx$$

- Αν είχαμε Gaussian κατανομή τότε το παραπάνω δηλώνεται:



- 90% διπλής – περιοχής διάστημα εμπιστοσύνης, α = 0.1
- 5% του ολοκληρώματος στις γραμμοσκιασμένες περιοχές σε κάθε πλευρά

Επίπεδα εμπιστοσύνης για δύο συνήθεις κατανομές

- Για Gaussian κατανομή:

  Περιοχή των άκρων α έξω από ±δ από τη μέση τιμή μιας Gaussian κατανομής:

α	δ	α	δ
0.3173	1σ	0.2	1.28σ
4.55×10^{-2}	2σ	0.1	1.64σ
2.7×10^{-3}	3σ	0.05	1.96σ
6.3×10^{-5}	4σ	0.01	2.58σ
5.7×10^{-7}	5σ	0.001	3.29σ
2.0×10^{-9}	6σ	10^{-4}	3.89σ

- Για Poissonian κατανομή:

  Χαμηλότερο και υψηλότερο (μονόπλευρη) όριο για την μέση τιμή *μ* μιας Poissonian μεταβλητής που δίνει *n* παρατηρούμενα γεγονότα απουσία υποβάθρου για διάστημα εμπιστοσύνης 90% και 95%:

	$1-\alpha$ =90%		$1-\alpha$ =95%	
n	$\mu_{\rm lo}$	$\mu_{\rm up}$	$\mu_{\rm lo}$	$\mu_{\rm up}$
0	–	2.30	–	3.00
1	0.105	3.89	0.051	4.74
2	0.532	5.32	0.355	6.30
3	1.10	6.68	0.818	7.75
4	1.74	7.99	1.37	9.15
5	2.43	9.27	1.97	10.51
6	3.15	10.53	2.61	11.84
7	3.89	11.77	3.29	13.15
8	4.66	12.99	3.98	14.43
9	5.43	14.21	4.70	15.71
10	6.22	15.41	5.43	16.96

- Στην περίπτωση αυτή το *α* αντιστοιχεί στο ποσοστό έξω από την περιοχή ολοκλήρωσης

Έλεγχος υπόθεσης (hypothesis testing)

- Ότι αναφέραμε μέχρι τώρα αποσκοπούσαν στην εύρεση της καλύτερης τιμής των παραμέτρων και την αβεβαιότητά τους κάτω από την υπόθεση ότι γνωρίζουμε την pdf :

- Τίποτα στη διαδικασία που ακολουθήσαμε δεν μας λέει αν τα δεδομένα που έχουμε συνάδουν με την υπόθεση που κάναμε

- Χρειαζόμαστε ένα στατιστικό test για να εξακριβώσουμε αν η υπόθεση που κάναμε είναι σωστή

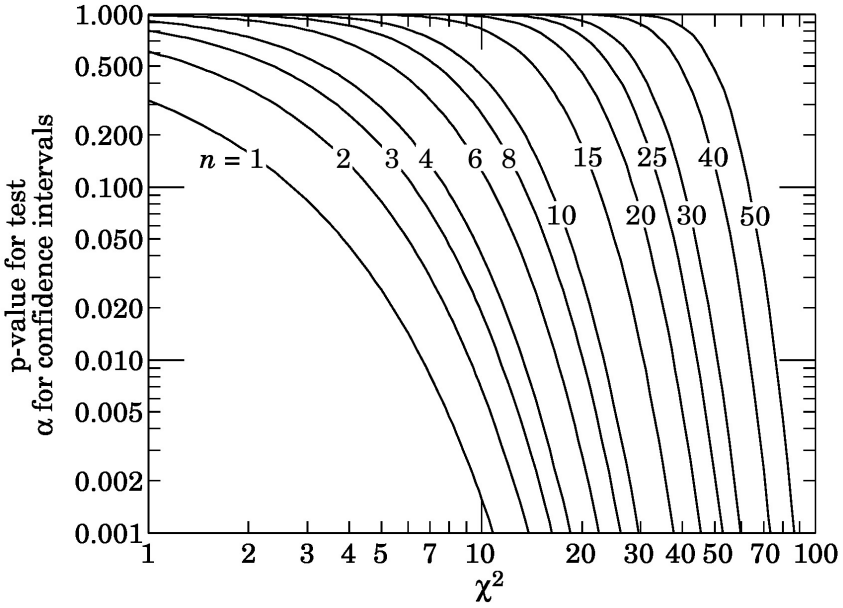
  - Τεστ σημαντικότητας: Πόσο πιθανόν είναι το σήμα να αποτελεί στατιστική διακύμανση

  - Τεστ καλής ποιότητας της προσαρμογής (Goodness of fit – GoF): έλεγχος για το κατά πόσο τα δεδομένα συνάδουν με την προτεινόμενη υπόθεση

  - Τεστ εξαίρεσης (exclusion test): ποιο το μέγεθος του σήματος το οποίο μπορεί να κρύβεται στα δεδομένα.

Τεστ σημαντικότητας

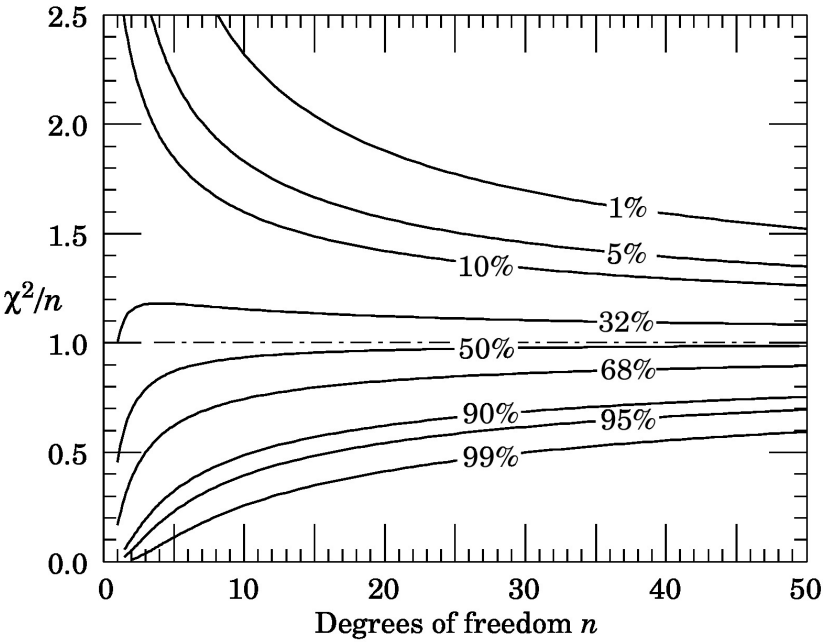
- Ας υποθέσουμε ότι μετρούμε μια τιμή *t* για τα δεδομένα:
  - Πόσο πιθανό είναι να δούμε μία τιμή η οποία είναι πιο μεγάλη από πρόβλεψη από την μέτρηση που κάναμε;



- Ας υποθέσουμε ότι μετρούμε μια κατανομή από δεδομένα
  - Πόσο συνάδει η κατανομή αυτή με την υπόθεση;

- Μπορούμε να χρησιμοποιήσουμε κάτι σαν χ^2

$$P-value=\int_{x^2_{meas.}}^{\infty} f\left(x;n_d\right)dx$$



Έλεγχος Υπόθεσης: Ο λόγος των πιθανοτήτων

- Στα περισσότερα πειράματα, όταν διερευνάται η ύπαρξη σήματος, αυτό βρίσκεται αναμεμειγμένο με αρκετό υπόβαθρο
- Πώς μπορούμε να ξέρουμε αν υπάρχει σημαντικό σήμα επιπλέον του υποβάθρου;
- Δεδομένων δύο υποθέσεων H_B (background) και H_{S+B} (σήμα και υπόβαθρο), ο λόγος των πιθανοτήτων αποτελεί ένα χρήσιμο στατιστικό τέστ

$$\lambda\left(\vec{N}\right) = \frac{\mathcal{L}\left(\vec{N}\middle| H_{S+B}\right)}{\mathcal{L}\left(\vec{N}\middle| H_{B}\right)}$$