Supreme Course I

지진원 특성평가
Characterization of Seismic Sources

2025.4.15~17

국립부경대학교 장보고관

PNSC PN's Solution

지진원 특성평가

Characterization of Seismic Sources

- Part I -

교육일정 및 내용

1일차	➢ 교육준비 ▪ 전산 프로그램 배포 및 설치 ➢ 교육과정 소개 ▪ 교육의 목표 및 내용 ➢ 기초 이론 ▪ 확률이론의 기초 ▪ 확률적 추정 (Probabilistic Estimation) ▪ 통계적 검정 (Statistical Test) ▪ 확률변수의 수치적 모사 (Monte Carlo Simulation)
2일차	➢ 지진목록 준비 ▪ 지진원 요소 ▪ 지진목록 병합 ➢ 지진목록의 완전성 평가 ▪ 배경 ▪ 완전성 평가방법의 분류 ▪ 지진목록을 이용한 완전성 평가 ➢ 지진규모 분포모델 ▪ 지수 모델 ▪ 특성지진 모델

교육일정 및 내용 (계속)

2일차 (계속)	➢ 지진원 특성평가 - 지진목록 이용 ▪ 지진원의 종류 및 요소 ▪ Richter-b 평가 ▪ 지진발생률 평가 ▪ 최대지진 평가 ▪ 반복적 동시평가
3일차	➢ 지질 및 측지자료의 이용 ▪ 최대지진 평가 ▪ 지진발생률 평가 ➢ 관련 이슈 ▪ 지진목록의 병합 효과 ▪ 지진의 이중 산입 ➢ SeisParEst를 이용한 실습 ▪ SeisParEst 사용자 지침 ▪ 지진원별 지진목록 작성: 지진원에 속하는 지진 추출 ▪ 지진목록의 완전성 평가: 6가지 방법 ▪ 지진원 특성 평가: 11가지 방법 ▪ 평가결과의 해석 및 활용
특전	동일 단체에서 2인 이상 수강하면, SeisParEst 1년 라이선스 제공

Chapter 0
Introduction

Preparation

▣ SeisParEst

- ❖ GUI-based computer code
- ❖ Construction of local catalogs
- ❖ Evaluation of catalog completeness
  - ➢ 6 methods
- ❖ Estimation of maximum potential earthquakes
  - ➢ 11 methods
- ❖ Estimation of a & b values
  - ➢ Linked together with m_{max} estimation

▣ Installation

- ❖ Copy SeisParEst.exe & SeisParEst.exe.manifest onto a same folder
- ❖ To run the program, double-click the SeisParEst.exe ()

About the Course

▣ Target Trainees

- ❖ Graduate/undergraduate students
- ❖ PSHA practitioners

▣ Goals

- ❖ To understand basic statistical seismology
- ❖ To evaluate seismicity parameters

▣ Contents

- ❖ Fundamental Statistics
- ❖ Construction & Assessment of local catalogs
- ❖ Estimation of seismicity parameters

Chapter 1
Fundamental Statistics

▣ Two Kinds of Probability Expression

❖ For two variables a and b belong to two sets A and B

- $a \in A$ and $b \in B$

❖ Intersection

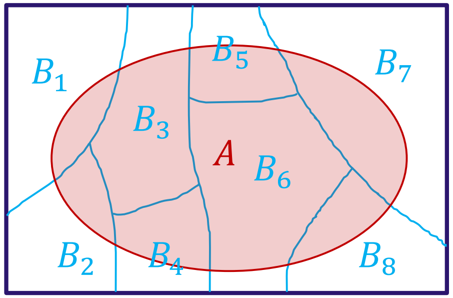
- $P(A \cap B) \leftrightarrow f_{AB}(a, b)$

❖ Independency

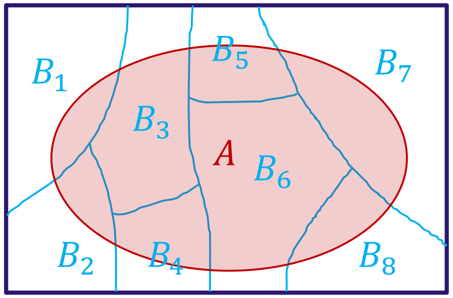
- $P(A \cap B) = P(A)P(B) \leftrightarrow f_{AB}(a, b) = f_A(a)f_B(b)$

♣ MECE principle

- Exclusiveness
  - $P(A \cap B) = \emptyset$
  - $P(A \cup B) = P(A) + P(B) - P(A \cap B) = P(A) + P(B)$
- Partition
  - If a subset $\{B_i\}$ of B is a partition of a union,
  - Mutually Exclusive (ME): $P\left(B_i \cap B_j\right) = \emptyset,\ if\ i \neq j$
  - Comprehensively Exhaustive (CE): $P(B_1 \cup B_2 \cdots \cup B_N) = \sum_i P(B_i) = 1$



▣ Two Kinds of Probability Expression (continued)



❖ Total probability

- If a subset $\{B_i\}$ of B is a partition of a union,
- $P(A) = \sum_i P(B_i \cap A) \leftrightarrow f_A(a) = \int_B f_{AB}(a,b)db$
- $f_A(a)$ is also called a marginal distribution

❖ Conditional probability

- $P(A|B) = P(A \cap B)/P(B) \leftrightarrow f_{A|B}(a|b) = f_{AB}(a,b)/f_B(b)$
- Since $f_{B|A}(b|a) = f_{AB}(a,b)/f_A(a)$
  - $f_{AB}(a,b) = f_{A|B}(a|b)f_B(b) = f_{B|A}(b|a)f_A(a)$

▣ Bayes' Theorem

❖ Bayes' rule

- $f_{M|D}(m|d) = \dfrac{f_{MD}(m,d)}{f_D(d)}$ conditional probability

  $= \dfrac{f_{MD}(m,d)}{\int f_{MD}(m,d)dm}$ total probability

  $= \dfrac{f_{D|M}(d|m)f_M(m)}{\int f_{D|M}(d|m)f_M(m)dm}$ conditional probability

  - $f_M(m)$: prior distribution or *a priori* information
  - $f_{D|M}(d|m)$: likelihood
  - $f_{M|D}(m|d)$: posterior distribution or update of $f_M(m)$

❖ Geophysical view point

- Conversion of the inverse problem into the forward problem
- If D is a set of observations and M the model parameters
  - $f_{M|D}(m|d)$: inversion of model parameter from observation
  - $f_{D|M}(d|m)$: forward calculation for a given set of model parameters
- To apply the Bayes' theorem, we need the distribution of model parameters, *a priori* information, which is not generally known

▣ Bayes' Theorem (continued)

❖ More comments

- Given a set of d, the numerator, $\int f_{D|M}(d|m)\, f_M(m)dm$ is a pure number
- Therefore, the following notation is frequently found
  - $f_{M|D}(m|d) = \tilde{f}_M(m) \propto f_{D|M}(d|m) f_M(m)$
  - $\tilde{f}_M(m)$, or equally $f_{M|D}(m|d)$ can be interpreted as a distribution of m improved by the observation d

❖ Bayesian distribution

- $\tilde{f}_D(d) = \int f_{DM}(d,m)dm = \int f_{D|M}(d|m)\, \tilde{f}_M(m)dm$

  $\leftrightarrow f_D(d) = \int f_{DM}(d,m)dm = \int f_{D|M}(d|m)\, f_M(m)dm$
- $\tilde{f}_D(d)$, the Bayesian distribution, can be interpreted as a weighted average of all possible density functions $\int f_{D|M}(d|m)$ which are associated with different values of M
- Here, the weight is the posterior distribution $\tilde{f}_M(m)$ which were improved or updated distribution by the Bayes' rule

▣ Bayes' Theorem (continued)

❖ Example 1: Simple application of the Bayes' rule (Cornell, 1972)

➢ Problem

- Reliability verification of a component which has never been designed, built, or tested before

➢ Assumption

- The failure of the component follows the Poisson process with the failure rate (number of failure per hour) of λ
  - Distribution of inter-failure time: $f_T(t) = \lambda e^{-\lambda t} \rightarrow P[T > t] = e^{-\lambda t}$
- $\lambda_1 = 0.001$ if the design team did successful job; $\lambda_2 = 0.01$ otherwise
- The reliability engineer knows, from his information on the design team (prior information), $P[\lambda = \lambda_1] = 0.9$ and $P[\lambda = \lambda_2] = 0.1$
- A single specimen has been tested for 300 hours ($= 1/\lambda$), then the test was terminated for economic reasons

➢ Evaluation

- The probability of observing a lifetime in excess of 300 hours is $P[T > 300] = e^{-300\lambda}$; call this event A then
- $P[\lambda = \lambda_1 | A] \propto P[A|\lambda = \lambda_1] \times P[\lambda = \lambda_1]$

$$\propto e^{-\frac{300}{1000}} \times 0.9 = 0.741 \times 0.9 = 0.247$$

▣ Bayes' Theorem (continued)

❖ Example 1: Simple application of the Bayes' rule (Continued)

- $P[\lambda = \lambda_2|A] \propto P[A|\lambda = \lambda_2] \times P[\lambda = \lambda_2]$

$$\propto e^{-\frac{300}{100}} \times 0.1 = 0.0498 \times 0.1 = 0.005$$

- The absolute values of these posterior probabilities are found by normalizing;
- $P[\lambda = \lambda_1|A] = \frac{0.247}{0.247+0.005} = 0.976 = \tilde{P}[\lambda = \lambda_1]$
- $P[\lambda = \lambda_2|A] = \frac{0.005}{0.247+0.005} = 0.024 = \tilde{P}[\lambda = \lambda_2]$

➢ Interpretation

- The prior information on the failure rate, $P[\lambda = \lambda_1] = 0.9$ and $P[\lambda = \lambda_2] = 0.1$, has been improved (updated) using the data from the 300 hour test
- The resultant posterior information says $\tilde{P}[\lambda = \lambda_1] = 0.976$ and $\tilde{P}[\lambda = \lambda_2] = 0.024$
- Note that, since we have only two cases, $\lambda = \lambda_1$ or $\lambda = \lambda_2$, the numerator of the Bayes' rule is

$$\begin{aligned} P[A] &= \sum_{i=1}^{2} P[A, \lambda_i] \\ &= P[A|\lambda = \lambda_1] \times P[\lambda = \lambda_1] + P[A|\lambda = \lambda_2] \times P[\lambda = \lambda_2] \\ &= 0.247 + 0.005 \end{aligned}$$

▣ Bayes' Theorem (continued)

❖ Example 2: The Bayes' distribution for the uncertain Richter-b

➢ Assumption

- Prior information: the Richter-b follows a gamma distribution
  - $f_{\mathrm{B}}(\beta) = k_1 \beta^{v-1} e^{-u\beta}$, where $k_1 = u^v / \Gamma(v)$ and $\beta = b \ln 10$
- Magnitudes follows a exponential distribution
  - $f_M(m) = \beta e^{-\beta(m-m_0)}$, $m \geq m_0$
- We have n observations of earthquake magnitude $[m_1, m_2, \cdots, m_n]$

➢ Task 1: Update $f_{\mathrm{B}}(\beta)$ using the observations of earthquakes

- $$\begin{aligned} l(sample|\beta) &= \beta e^{-\beta(m_1-m_0)} \beta e^{-\beta(m_1-m_0)} \ldots \beta e^{-\beta(m_n-m_0)} \\ &= \beta^n \exp[-\textstyle\sum_{i=1}^n \beta\,(m_i - m_0)] \\ &= \beta^n \exp[-n\beta(\bar{m} - m_0)\,] \qquad \because \bar{m} = \textstyle\sum_{i=1}^n m_i \\ &= \beta^n \exp(-n\beta\hat{m}) \qquad \because \hat{m} = \bar{m} - m_0 \end{aligned}$$
- $$\begin{aligned} \tilde{f}_{\mathrm{B}}(\beta) &\propto l(sample|\beta) f_{\mathrm{B}}(\beta) \\ &\propto \beta^n \exp(-n\beta\hat{m})\, \beta^{v-1} e^{-u\beta} \\ &= k_2 \beta^{n+v-1} \exp[-\beta(n\hat{m} + u)] \\ &= k_2 \beta^{v'-1} e^{-u'\beta} \end{aligned}$$ (Cornell, 1972; Campbell, 1982)

  where $v' = n + v$, $u' = n\hat{m} + u$, and $k_2 = (u')^{v'} / \Gamma(v')$
- Updated distribution, $\tilde{f}_{\mathrm{B}}(\beta)$, is again a gamma distribution

Bayes' Theorem (continued)

❖ Example 2: Uncertain Richter-b (continued)

- In the distribution, $f_{\mathrm{B}}(\beta) = k_1 \beta^{v-1} e^{-u\beta}$, the mean and variance of β are $\bar{\beta} = v/u$ and $\sigma_\beta^2 = v/u^2$ which can be interpreted as the prior 'best estimates' of the mean and variance of β
- Using these relations, we have: $v' = n + \left(\frac{\bar{\beta}}{\sigma_\beta}\right)^2$ and $u' = n(\bar{m} - m_0) + \frac{\bar{\beta}}{\sigma_\beta^2}$

➤ Task 2: Update $f_M(m)$ to get the Bayesian distribution, using $\tilde{f}_{\mathrm{B}}(\beta)$

- Starting with $m_{max} = \infty$, the updated distribution is given by
- $\tilde{F}_M(m) = \int_0^\infty F_M(m|\beta) \tilde{f}_{\mathrm{B}}(\beta) d\beta$

$$= \int_0^\infty \left[1 - e^{-\beta(m-m_0)}\right] k_2 \beta^{v'-1} e^{-u'\beta} d\beta$$

$$= 1 - k_2 \int_0^\infty \beta^{v'-1} e^{-u''\beta} d\beta \qquad \because u'' = u' + m - m_0$$

$$= 1 - k_2 \frac{\Gamma(v')}{(u'')^{v'}} = 1 - \left(\frac{u'}{u''}\right)^{v'}$$

$$= 1 - \left(\frac{u'}{u'+m-m_0}\right)^{v'}, \quad m_0 \le m < \infty$$

(Campbell, 1982)

▣ Bayes' Theorem (continued)

❖ Example 2: Uncertain Richter-b (continued)

- Introducing the maximum magnitude, m_{max} and the normalization constant, K

$$K[\tilde{F}_M(m_{max}) - \tilde{F}_M(m_0)] = 1 \text{ or } K = \left[1 - \left(\frac{u'}{u'+m_{max}-m_0}\right)^{v'}\right]^{-1}$$

- $\tilde{F}_M(m) = \begin{cases} 0, & m < m_0 \\ K\left[1 - \left(\frac{u'}{u'+m-m_0}\right)^{v'}\right], & m_0 \le m \le m_{max} \\ 1, & m > m_{max} \end{cases}$ (Campbell, 1982)

where $\begin{cases} v' = n + v = n + \left(\frac{\bar{\beta}}{\sigma_\beta}\right)^2 \\ u' = n\hat{m} + u = n(\bar{m} - m_0) + \frac{\bar{\beta}}{\sigma_\beta^2} \end{cases}$

▣ Characterization of Distributions

❖ Notation

➢ Random variables are denoted by capital letters such as X while the values taken by random variables by lowercase letters such as x

❖ Probability density function (PDF)

➢ $P(x \le X \le x + dx) = f_X(x)dx, \quad x \in [a, b]$

➢ $f_X(x) = \begin{cases} \frac{P(x \le X \le x+dx)}{dx}, & [a, b] \\ 0, & \text{otherwise} \end{cases}$

❖ Cumulative distribution function (CDF)

➢ $F_X(x) = P(X \le x) = \int_{-\infty}^{x} f_X(x)dx$

$$= \int_a^x f_X(x)dx \leftrightarrow f_X(x) = \frac{dF_X(x)}{dx}$$

➢ $F_X(x) = \begin{cases} 0, & x \le a \\ \int_a^x f_X(x)dx, & a \le x \le b \\ 1, & x > b \end{cases}$

▣ Representative Values

❖ Location

➢ Mode

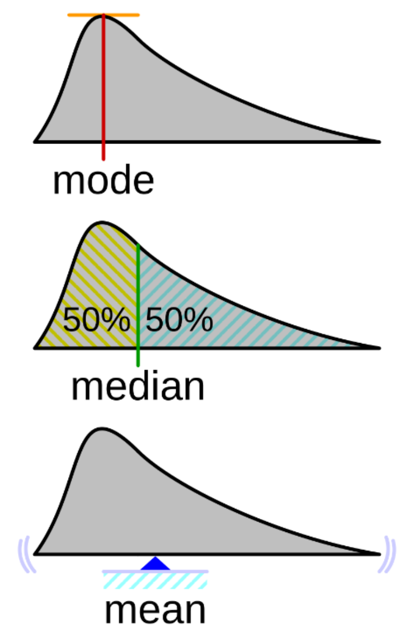
- A value that most frequently occurs

➢ Median (50th percentile)

- A value separating the higher half from the lower half of a data sample, a population, or a probability distribution

➢ Mean (expectation)

- For the discrete random variable: $E(X) = \sum_i p_i x_i$
- For the continuous random variable: $E(X) = \int x f_X(x) dx$
- Linear operator:
  - $E[a \cdot g(X) + b \cdot h(Y)] = a \int g(x) f_X(x) dx + b \int h(y) f_Y(y) dy$
    $= aE[g(X)] + bE[h(Y)]$
  - $E[aX + b] = aE[X] + b$



<from Wikipedia>

Representative Values (continued)

Scale

Variance

- $Var(X) = E[(X-\mu)^2] = E[X^2 - 2\mu X + \mu^2]$
  $= E[X^2] - 2\mu E[2X] + \mu^2$
  $= E[X^2] - \mu^2$, where $\mu = E[X]$
- $Var(aX+b) = E\{[(aX+b)-(a\mu-b)]^2\}$
  $= E[a^2(X-\mu)^2]$
  $= a^2 E[(X-\mu)^2]$
  $= a^2 Var(X)$

Standard deviation

- $\sigma(X) = \sqrt{Var(X)}$
- $\sigma(aX+b) = |a|\sqrt{Var(X)}$

Quantiles

Definition

- A quantile is a cut point that divides a probability distribution's range into continuous intervals

Percentile

- A cut point that divides a probability distribution's range into 100 equal continuous intervals

Decile

- A cut point that divides a probability distribution's range into 10 equal continuous intervals

Quartile

- A cut point that divides a probability distribution's range into 4 equal continuous intervals
- Interquartile range (IQR)
  - $IQR = x_{0.75} - x_{0.25}$ → range including a half of data
  - For Gaussian distribution, $IQR = 1.349\sigma$
    - Pseudo-standard deviation: $S_{ps} = IQR/1.349$

♣Resistance & Robustness

- Resistance
  - Degree of tolerance of a statistical technique (an estimator or a statistical test) to the presence of outliers
  - Ex: median has the maximum resistance of 0.5
- Robustness
  - Insensitivity with regard to an underlying assumed probability model
  - Ex: residuals are assumed to follow a Gaussian or a uniform distribution with zero mean

▣ Correlations

❖ Covariance

➢ $Cov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)]$

$$= E[XY - \mu_X Y - \mu_Y X + \mu_X \mu_Y]$$

$$= E[XY] - \mu_X \mu_Y$$

➢ $Cov(X,Y) = 0$, if X and Y are independent

➢ $Cov(aX+b, cY+d) = E[a(X-\mu_X)c(Y-\mu_Y)]$

$$= ac\, Cov(X,Y)$$

❖ Correlation Coefficient

➢ $Corr(X,Y) = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}, \quad -1 \le Corr(X,Y) \le +1$

➢ $Corr(aX+b, cY+d) = \frac{ac\, Cov(X,Y)}{|a||c|\sigma_X \sigma_Y} = \mathrm{sgn}(a)\,\mathrm{sgn}(c) Corr(X,Y)$

♣ Coefficient of variation

➢ $CV(X) = \frac{\sigma}{\mu}$

➢ Frequently denoted by CoV

Sample Mean & Variance

▣ Random Sample

For $X_1, X_2, \cdots, X_n$ sampled from a population with mean μ and variance σ^2

- ❖ Each sample X_i is a random variable
- ❖ Value x_i of a sample X_i is a realization of X_i
- ❖ The set $\{X_1, X_2, \cdots, X_n\}$ is called a random sample of X, of which size is n

▣ Statistic

- ❖ A function of random sample
- ❖ Since a random sample is the set of random variables, a statistic is a random variable also

▣ Sample mean

❖ For X_i sampled from a population with a mean μ and variance σ^2

❖ Definition: $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

❖ Mean of sample mean:

➢ $E(\bar{X}) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \mu$ ($\bar{X}$ is an unbiased estimator of μ)

❖ Variance of sample mean:

➢ $E(\bar{X}^2) = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} E(X_i X_j)$

$$= \frac{1}{n^2}[n(n-1)\mu^2 + n(\sigma^2 + \mu^2)] = \mu^2 + \frac{\sigma^2}{n}$$

$$E(X_i X_j) = \begin{cases} E(X_i)E(X_j) = \mu^2, & i \neq j \\ E(X_i^2) = \sigma^2 + \mu^2, & i = j \end{cases}$$

➢ $Var(\bar{X}) = E(\bar{X}^2) - E^2(\bar{X})$

$$= \left(\mu^2 + \frac{\sigma^2}{n}\right) - \mu^2 = \frac{\sigma^2}{n}$$

♣ For a large, n from the central limit theorem, $\bar{X} \sim N(\mu, \sigma^2/n)$

▣ Sample variance

❖ Definition:

➢ $$V = \begin{cases} \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2, & \text{for unknown } \mu \\ \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2, & \text{for known } \mu \end{cases}$$

❖ Mean of sample variance

➢ For unknown μ

$$E(V) = \frac{1}{n-1}\sum_{i=1}^{n}\left[E\left(X_i^2\right) - 2E(X_i\bar{X}) + E(\bar{X}^2)\right]$$

$$= \frac{1}{n-1}\sum_{i=1}^{n}\left[(\sigma^2 + \mu^2) - 2\left(\mu^2 + \frac{\sigma^2}{n}\right) + (\mu^2 + \frac{\sigma^2}{n})\right]$$

$$= \frac{1}{n-1}\sum_{i=1}^{n}\left[\frac{n-1}{n}\sigma^2\right] = \sigma^2$$

- unbiased estimator
- degrees of freedom decreased by 1

❖ Task: show that $E(X_i\bar{X}) = \mu^2 + \frac{\sigma^2}{n}$

▣ Sample variance (continued)

❖ Mean of sample variance (continued)

➢ For known μ

$$E(V) = \frac{1}{n}\sum_{i=1}^{n}\left[E\left(X_i^2\right) - 2\mu E(X_i) + \mu^2\right]$$

$$= \frac{1}{n}\sum_{i=1}^{n}[(\sigma^2 + \mu^2) - \mu^2]$$

$$= \sigma^2$$

- unbiased estimator

Binomial Distribution

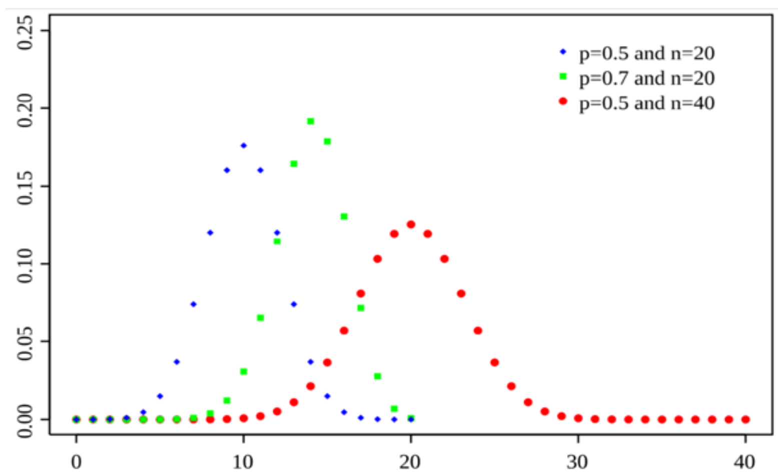
- Bernoulli Trial
  - $S = \{s, f\}$
  - $p = P\{s\} \geq 0,\ q = P\{f\} \geq 0;\ \ p + q = 1$

- Binomial distribution, $B(n, p)$
  - X : frequency of success in the n independent Bernoulli trials
    - $P\{X = x\} = \binom{n}{x} p^x q^{n-x},\ x = 0, 1, \cdots, n$
  - The whole distribution can be expressed by binomial expansion

$$(p+q)^n = \sum_{x=0}^{n} \binom{n}{x} p^x q^{n-x}$$

  - Mean:

$$\begin{aligned} E(X) &= \sum_{x=0}^{n} x \binom{n}{x} p^x q^{n-x} \\ &= p \frac{\partial}{\partial p} \left[\sum_{x=0}^{n} \binom{n}{x} p^x q^{n-x} \right] \\ &= p \frac{\partial}{\partial p} (p+q)^n = np(p+q)^{n-1} \\ &= np \qquad \because p + q = 1 \end{aligned}$$



<from Wikipedia>

▣ Binomial Distribution (continued)

❖ Binomial distribution (continued)

➢ Variance:

$$E(X^2) = \sum_{x=0}^{n} x^2 \binom{n}{x} p^x q^{n-x}$$

$$= p\frac{\partial}{\partial p}\left\{p\frac{\partial}{\partial p}\left[\sum_{x=0}^{n}\binom{n}{x} p^x q^{n-x}\right]\right\}$$

$$= p\frac{\partial}{\partial p}\left[p\frac{\partial}{\partial p}(p+q)^n\right]$$

$$= np(p+q)^{n-1}+n(n-1)p^2\,(p+q)^{n-2}$$

$$= np + n(n-1)p^2$$

$$Var(X) = E(X^2) - E^2(X)$$

$$= [np + n(n-1)p^2] - (np)^2$$

$$= np(1-p) = npq$$

❖ Sum of binomial deviates

➢ If X_1 and X_2 are mutually independent, and $X_1 \sim B(n, p)$ and $X_2 \sim B(m, p)$, then $X_1 + X_2 \sim B(n + m, p)$

▣ Poisson Distribution

❖ Poisson process

- For non-overlapping unit intervals, the occurrence frequency in one unit interval is independent of that in another (independent, memoryless)
- The probability of more than one occurrence in an extremely small interval is extremely small
- The mean occurrence frequency in a unit interval is constant and time-invariant: *homogeneous* Poisson process

❖ Derivation of distribution from binomial distribution

- For large n with $m = np$

$$P\{X=x\} = \binom{n}{x} p^x q^{n-x}$$

$$= \frac{1}{x!} n(n-1)\cdots(n-x+1)\left(\frac{m}{n}\right)^x \left(1-\frac{m}{n}\right)^{n-x}$$

$$= \frac{m^x}{x!}\left[1\left(1-\frac{1}{n}\right)\cdots\left(1-\frac{x-1}{n}\right)\right]\left(1-\frac{m}{n}\right)^n \left(1-\frac{m}{n}\right)^{-x}$$

$$\approx \frac{e^{-m}m^x}{x!} \quad \because \left(1-\frac{m}{n}\right)^n \approx e^{-m}$$

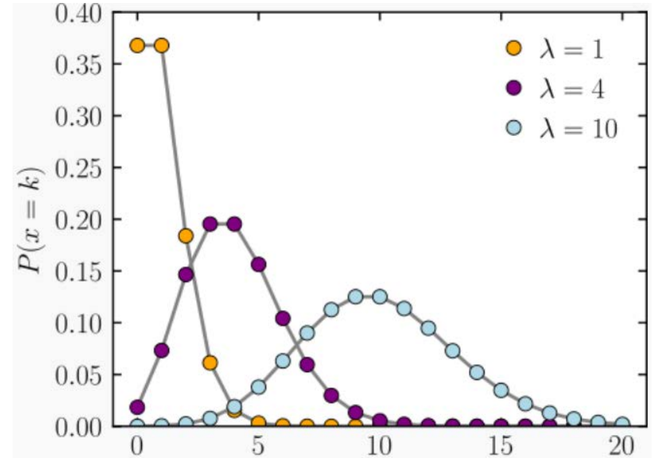
▣ Poisson Distribution (continued)

❖ Mean:

$$E(X) = \sum_{x=0}^{\infty} x \frac{e^{-m} m^x}{x!}$$

$$= m e^{-m} \frac{\partial}{\partial m}\left[\sum_{x=0}^{\infty} \frac{m^x}{x!}\right]$$

$$= m e^{-m} \frac{\partial}{\partial m}(e^m) = m$$



<PMF with $\lambda \equiv m$, from Wikipedia>

❖ Variance:

$$E(X^2) = \sum_{x=0}^{\infty} x^2 \frac{e^{-m} m^x}{x!}$$

$$= m e^{-m} \frac{\partial}{\partial m}\left[m \frac{\partial}{\partial m}\left(\sum_{x=0}^{\infty} \frac{m^x}{x!}\right)\right] = m e^{-m} \frac{\partial}{\partial m}(m e^m)$$

$$= m e^{-m}(e^m + m e^m) = m(1+m)$$

$$Var(X) = E(X^2) - E^2(X)$$

$$= m(1+m) - (m)^2 = m$$

▣ Poisson Distribution (continued)

❖ Inter-event time

- If λ is the rate, i.e., the frequency in unit time, the mean expectation of events during time t is $m = \lambda t$
- The probability for $X = x$ events is
  - $P\{X = x; m = \lambda t\} = \frac{e^{-\lambda t}(\lambda t)^n}{n!}$
- No event up to time time τ from the last event means that the inter-event time is larger than τ so that
  - $P\{X = 0; m = \lambda\tau\} = e^{-\lambda\tau} = 1 - F(\tau;\lambda)$ ← exponential distribution

❖ Sum of Poisson deviates

- If X_1 and X_2 are mutually independent, and $X_1 \sim P_X(m_1)$ and $X_2 \sim P_X(m_2)$, then $X_1 + X_2 \sim P_X(m_1 + m_2)$

▣ Exponential Distribution

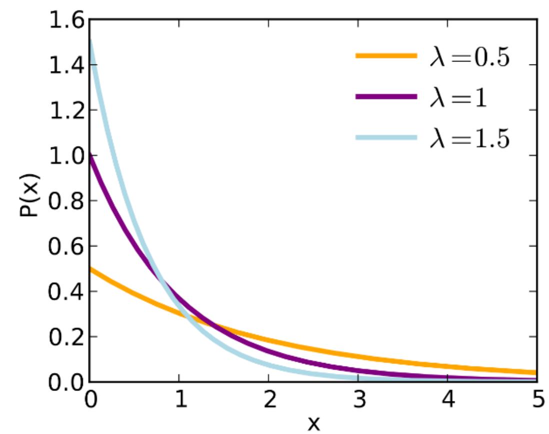
❖ PDF

- For rate parameter $\lambda > 0$, $f(x;\lambda) = \lambda e^{-\lambda x}, \quad x \in [0, +\infty)$
  - Mean: $1/\lambda$
  - Variance: $1/\lambda^2$
- $P(X > x) = 1 - F(x;\lambda) = e^{-\lambda x}$
- Memoryless: $P(X > s + x | X > s) = P(X > x)$ ≑ Poisson process

❖ Sum of exponential deviates

- If X_1 and X_2 are mutually independent exponential deviates with rates λ_1 and λ_2, respectively, then the PDF of $Z = X_1 + X_2$ is

$$f_Z(z) = \begin{cases} \frac{\lambda_1 \lambda_2}{\lambda_2 - \lambda_1}\left(e^{-\lambda_1 z} - e^{-\lambda_2 z}\right), & \lambda_1 \neq \lambda_2 \\ \lambda z e^{-\lambda z}, & \lambda_1 = \lambda_2 \end{cases}$$



<from Wikipedia>

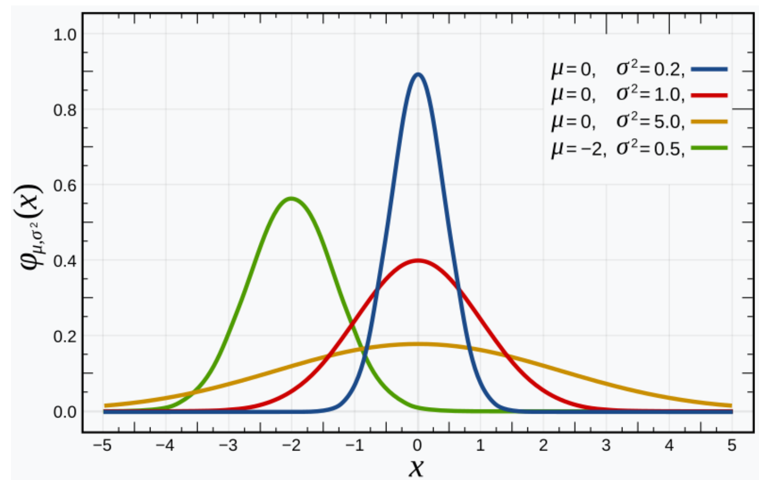
▣ Normal Distribution (Gaussian Distribution)

❖ Notation

➢ If a random variable follows the normal distribution with a mean μ and a variance σ^2, it is denoted by $X \sim N(\mu, \sigma^2)$

❖ Probability density function

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty$$



The red curve is the *standard normal distribution.*

<from Wikipedia>

❖ Mean: μ

$$0 = \frac{\partial}{\partial\mu}\int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^{\infty}\frac{\partial}{\partial\mu} f(x)dx$$

$$= \frac{1}{\sqrt{2\pi}\,\sigma}\left(\frac{1}{\sigma^2}\right)\int_{-\infty}^{\infty}(x-\mu)e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx$$

$$= \left(\frac{1}{\sigma^2}\right)[E(X)-\mu] \quad \therefore E(X) = \mu$$

❖ Variance: σ^2

$$0 = \frac{\partial^2}{\partial\mu^2}\int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^{\infty}\frac{\partial^2}{\partial\mu^2} f(x)dx = \frac{1}{\sqrt{2\pi}\,\sigma}\left(\frac{1}{\sigma^2}\right)\int_{-\infty}^{\infty}\left[\frac{1}{\sigma^2}(x-\mu)^2 - 1\right] e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx$$

$$= \left(\frac{1}{\sigma^2}\right)\left[\frac{1}{\sigma^2}Var(X) - 1\right] \quad \therefore Var(X) = \sigma^2$$

▣ Normal Distribution (continued)

❖ Standard normal distribution

- $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$

❖ Sum of normal deviates

- If X_1 and X_2 are mutually independent, and $X_1 \sim N(\mu_1, {\sigma_1}^2)$ and $X_2 \sim N(\mu_2, {\sigma_2}^2)$, then

$$X_1 \pm X_2 \sim N(\mu_1 \pm \mu_2, {\sigma_1}^2 + {\sigma_2}^2)$$

❖ Log-normal distribution

- $\log X \sim N(\mu_{ln}, \sigma_{ln}^2)$

▣ Gamma Distribution

❖ Gamma function

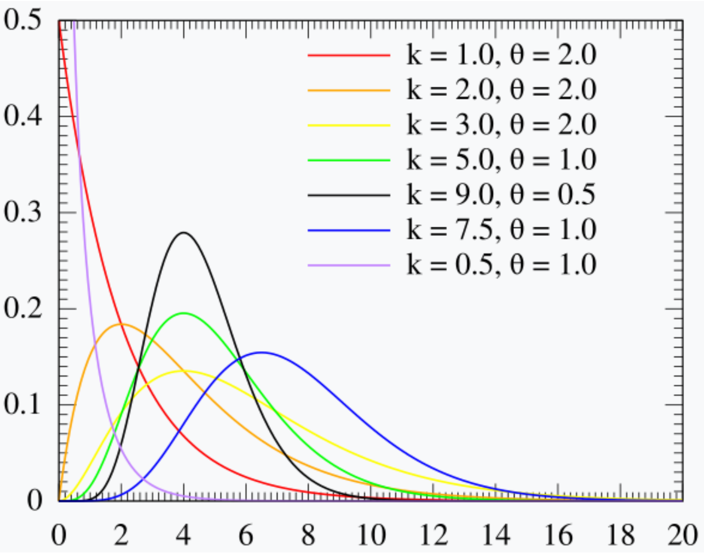
➢ Complete gamma function

- $\Gamma(b) = \int_0^\infty z^{b-1} e^{-z} dz, \quad b > 0$
- $\Gamma(b+1) = b\Gamma(b)$

➢ Incomplete gamma functions

- Upper: $\Gamma(x; b) = \int_x^\infty z^{b-1} e^{-z} dz$
- Lower: $\gamma(x; b) = \int_0^x z^{b-1} e^{-z} dz$

➢ Note that several different notations are in use



<Probability density function, $\mathrm{k} \equiv b, \ \theta \equiv 1/c$>
<from Wikipedia>

❖ Probability density function

$$f(x) = ax^{b-1}e^{-cx}, \quad x > 0, \quad b, c > 0$$

➢ Normalization:

- $1 = a \int_0^\infty x^{b-1} e^{-cx} dx$

$$= \frac{a}{c} \int_0^\infty (w/c)^{b-1} e^{-w} dw \qquad \because w = cx$$

$$= ac^{-b} \int_0^\infty w^{b-1} e^{-w} dw = ac^{-b}\Gamma(b) \qquad \therefore a = c^b/\Gamma(b)$$

Gamma Distribution (continued)

- Mean: $\mu = \frac{a}{c^b}\frac{\Gamma(b+1)}{c} = \frac{b}{c}$
- Variance: $\sigma^2 = \frac{b}{c^2}$
- Complete notation: $f(x) \rightarrow f(x; b, c)$

❖ Cumulative distribution function

- $F(x; b, c) = \int_0^x f(z; b, c)dz$

$$= a\int_0^x z^{b-1}e^{-cz}dz$$

$$= \frac{\gamma(cx; b)}{\Gamma(b)}$$

❖ Notation: $x \sim G(b, c)$

- b is called a shape parameter and c a rate parameter

❖ Sum of gamma deviates

- If X_1 and X_2 are mutually independent, and $X_1 \sim G(b_1, c)$ and $X_2 \sim G(b_2, c)$, then $X_1 + X_2 \sim G(b_1 + b_2, c)$

♣ Normal & Gamma Distributions

❖ Moment of normal distribution

➢ $G(x;\, n) = \frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{x} z^n e^{-\frac{(z-\mu)^2}{2\sigma^2}} dz$

$$= \frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{\frac{z-\mu}{\sigma}} (\sigma w + \mu)^n e^{-\frac{w^2}{2}} (\sigma dw) \qquad \because w = \frac{z-\mu}{\sigma}$$

$$= \frac{1}{\sqrt{2\pi}}\sum_{i=0}^{n}\binom{n}{i}\sigma^i \mu^{n-i}\int_{-\infty}^{\frac{z-\mu}{\sigma}} w^i e^{-\frac{w^2}{2}} dw$$

$$= \frac{1}{\sqrt{2\pi}}\sum_{i=0}^{n}\binom{n}{i}\sigma^i \mu^{n-i}\begin{cases}\int_{-\infty}^{0} w^i e^{-\frac{w^2}{2}} dw + \int_{0}^{\frac{z-\mu}{\sigma}} w^i e^{-\frac{w^2}{2}} dw & x \geq \mu \\ \int_{-\infty}^{0} w^i e^{-\frac{w^2}{2}} dw - \int_{\frac{z-\mu}{\sigma}}^{0} w^i e^{-\frac{w^2}{2}} dw & x < \mu\end{cases}$$

$$= \frac{1}{\sqrt{2\pi}}\sum_{i=0}^{n}\binom{n}{i}\sigma^i \mu^{n-i}\begin{cases} I_0(i) + I_+(x; i) & x \geq \mu \\ I_0(i) - I_-(x; i) & x < \mu\end{cases}$$

♣ Normal & Gamma Distributions

❖ General formulation (continued)

- $I_0\,(i) = \int_{-\infty}^{0} w^i e^{-\frac{w^2}{2}} dw = (-1)^i \left(\sqrt{2}\right)^{i-1} \int_0^{\infty} v^{\frac{i-1}{2}} e^{-v} dv \quad \because v = \frac{w^2}{2}$

  $= (-1)^i \left(\sqrt{2}\right)^{i-1} \Gamma(\frac{i+1}{2})$

- $I_+\,(x;i) = \int_0^{(x-\mu)/\sigma} w^i e^{-\frac{w^2}{2}} dw = \left(\sqrt{2}\right)^{i-1} \int_0^{\frac{(x-\mu)^2}{2\sigma^2}} v^{\frac{i-1}{2}} e^{-v} dv\,, \ \ x \geq \mu$

  $= \left(\sqrt{2}\right)^{i-1} \gamma(\frac{(x-\mu)^2}{2\sigma^2}; \frac{i+1}{2})$

- $I_-\,(x;i) = \int_{(x-\mu)/\sigma}^{0} w^i e^{-\frac{w^2}{2}} dw = (-1)^i \left(\sqrt{2}\right)^{i-1} \int_0^{\frac{(x-\mu)^2}{2\sigma^2}} v^{\frac{i-1}{2}} e^{-v} dv\,, \ \ x < \mu$

  $= (-1)^i \left(\sqrt{2}\right)^{i-1} \gamma(\frac{(x-\mu)^2}{2\sigma^2}; \frac{i+1}{2})\,, \ \ x < \mu$

♣ Normal & Gamma Distributions (continued)

❖ Cumulative distribution

➢ $F(x) = G(x; n = 0)$

$$= \frac{1}{\sqrt{2\pi}} \begin{cases} I_0(0) + I_+(x;0) & x \geq \mu \\ I_0(0) - I_-(x;0) & x < \mu \end{cases}$$

$$= \frac{1}{2\sqrt{\pi}} \begin{cases} \sqrt{\pi} + \gamma(\frac{(x-\mu)^2}{2\sigma^2}; \frac{1}{2}) & x \geq \mu \\ \sqrt{\pi} - \gamma(\frac{(x-\mu)^2}{2\sigma^2}; \frac{1}{2}) & x < \mu \end{cases}$$

➢ Since $\gamma\left(\frac{(x-\mu)^2}{2\sigma^2}; \frac{1}{2}\right)\Big|_{x=\pm\infty} = \gamma\left(\infty; \frac{1}{2}\right) = \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$,

- $F(-\infty) = 0$ and $F(\infty) = 1$

❖ Mean

➢ $E(x) = G(\infty; n = 1)$ $\qquad \because \gamma(\infty; i) = \Gamma(i)\ \&\ I_+(\infty; i) = (-1)^i I_0(i)$

$$= \frac{1}{\sqrt{2\pi}} \{\mu[I_0\,(0) + I_+(\infty;0)] + \sigma[I_0\,(1) + I_+(\infty;1)]\}$$

$$= \frac{1}{\sqrt{2\pi}} \{2\mu I_0\,(0)\} = \frac{1}{\sqrt{2\pi}} \left\{2\mu \frac{\Gamma(1/2)}{\sqrt{2}}\right\} = \mu$$

♣ Normal & Gamma Distributions (continued)

❖ Variance

- $E(x^2) = G(\infty; n = 2)$

$$= \frac{1}{\sqrt{2\pi}}\{\mu^2[I_0(0) + I_+(\infty;0)] + 2\mu\sigma[I_0(1) + I_+(\infty;1)] + \sigma^2[I_0(2) + I_+(\infty;2)]\}$$

$$= \frac{1}{\sqrt{2\pi}}\{2\mu^2 I_0(0) + 2\sigma^2 I_0(2)\}$$

$$= \frac{1}{\sqrt{2\pi}}\left\{2\mu^2\frac{\Gamma(1/2)}{\sqrt{2}} + 2\sigma^2\left[\frac{\sqrt{2}\Gamma\left(\frac{3}{2}\right)}{2}\right]\right\}$$

$$= \mu^2 + \sigma^2$$

- $Var(x) = E(x^2) - E^2(x) = \sigma^2$

▣ χ^2 Distribution

❖ Chi-square deviate: $X = \sum_{i=1}^{k} Z_i^2 \sim \chi^2(k)$

➢ $Z_k \sim N(0,1)$ and k is degrees of freedom

❖ PDF

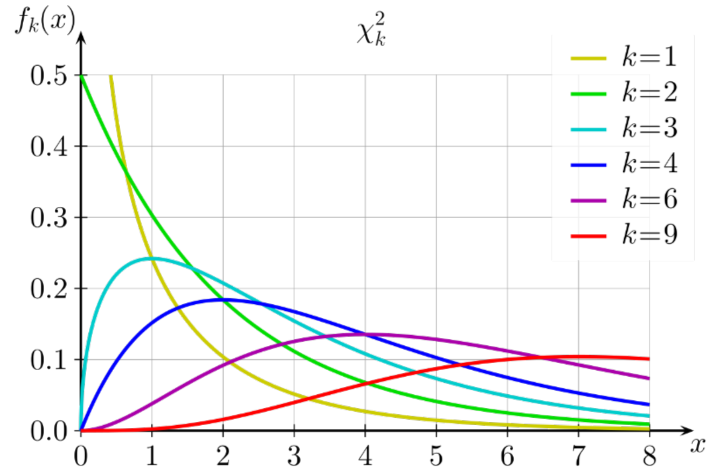
➢ $f(x;k) = \frac{x^{k/2-1}e^{-x/2}}{2^{k/2}\Gamma(k/2)}, \quad x \in [0,+\infty)$

➢ Mean: k

➢ Variance: $2k$

❖ CDF

➢ $F(x;k) = \frac{\gamma(x/2;k/2)}{\Gamma(k/2)}$



<from Wikipedia>

❖ Sum of χ^2 deviates

➢ If V_1 and V_2 are mutually independent, and $V_1 \sim \chi^2(k_1)$ and $V_2 \sim \chi^2(k_2)$, then $V_1 + V_2 \sim \chi^2(k_1 + k_2)$

▣ Student t Distribution

❖ Student t deviate: $T = \frac{Z}{\sqrt{V/\nu}}$

➢ $Z \sim N(0,1)$

➢ $V \sim \chi^2(\nu)$

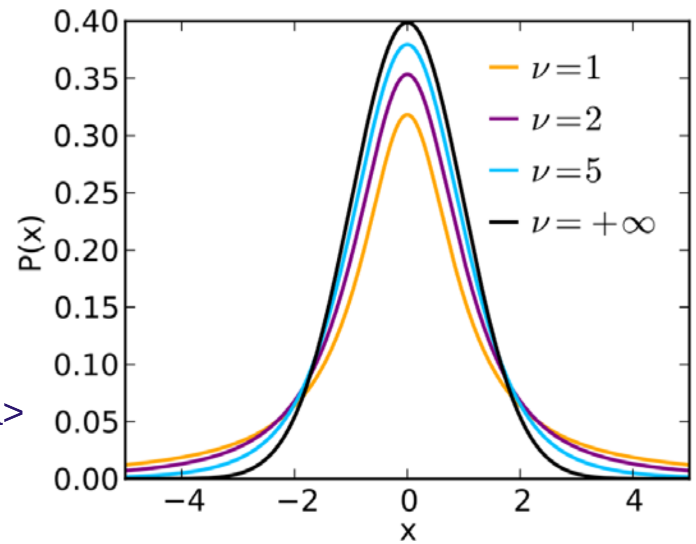
➢ PDF: $f(t;\nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu}\,\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}}$

- Mean: 0 for $\nu > 1$, otherwise undefined
- Variance: $\frac{\nu}{\nu-2}$ for $\nu > 2$; ∞ for $2 < \nu \leq 4$; otherwise undefined

❖ Usage

➢ To test a location of distribution

<from Wikipedia>



F Distribution

❖ Definition

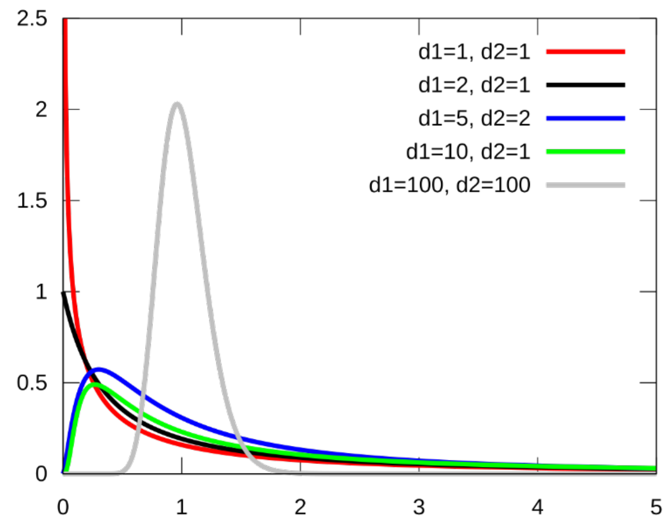
- F deviate: $F = \frac{V_1/\nu_1}{V_2/\nu_2} \sim F(\nu_1, \nu_2)$
  - $V_1 \sim \chi^2(\nu_1)$
  - $V_2 \sim \chi^2(\nu_2)$
- PDF: $F(x; \nu_1, \nu_2) = \frac{1}{xF\left(\frac{\nu_1}{2}, \frac{\nu_2}{2}\right)} \sqrt{\frac{(\nu_1 x)^{\nu_1} \nu_2^{\nu_2}}{(\nu_1 x + \nu_2)^{\nu_1 + \nu_2}}}$
  - Mean: $\frac{\nu_2}{\nu_2 - 2}$ for $\nu_2 > 2$
  - Variance: $\frac{2\nu_2^2(\nu_1+\nu_2-2)}{\nu_1(\nu_2-2)^2(\nu_2-4)}$ for $\nu_2 > 4$

❖ Useful properties

- $1/F = \frac{V_2/\nu_2}{V_1/\nu_1} \sim F(\nu_2, \nu_1)$
- $T^2 = \frac{Z^2}{V/\nu} \sim F(1, \nu)$

❖ Usage

- To test a variance of distribution



<from Wikipedia>

Order Statistics

Distributions of maxima

- Suppose that we have a set of n random X_i which have a PDF of $f_X(x)$
- If Y_k take the ordered values of X_i such that $Y_1 \le Y_2 \cdots \le Y_k \cdots \le Y_n$, then $Y_k = X_i \sim f_X(x)$
- Distribution functions
  - $Y = Y_n = max\{X_i\}$
  - $F_Y(y) = P(Y \le y)$

$$= P(X_1 \le y, X_2 \le y, \cdots, X_n \le y)$$

$$= P(X_1 \le y)\, P(X_2 \le y) \cdots P(X_n \le y)$$ if X_i are mutually independent

$$= F_{X_1}(y)\, F_{X_2}(y) \cdots F_{X_n}(y)$$

$$= [F_X(y)]^n$$ if X_i are identically distributed

  - $f_Y(y) = n f_X(y)[F_X(y)]^{n-1}$

▣ Order Statistics (continued)

❖ Distribution of minima

- $Z = Y_1 = min\{X_i\}$
- $$\begin{aligned} F_Z(z) = P(Z \le z) &= 1 - P(Z \ge z) \\ &= 1 - P(Y_1 \ge z, Y_2 \ge z, \cdots, Y_n \ge z) \\ &= 1 - P(X_1 \ge z, X_2 \ge z, \cdots, X_n \ge z) \\ &= 1 - P(X_1 \ge z)\, P(X_2 \ge z) \cdots P(X_n \ge z) \quad \text{mutually independent } X_i \\ &= 1 - \left[1 - F_{X_1}(z)\right]\left[1 - F_{X_2}(z)\right] \cdots \left[1 - F_{X_n}(z)\right] \\ &= 1 - [1 - F_X(z)]^n \quad \text{identically distributed } X_i \end{aligned}$$
- $f_Z(z) = n f_X(z)[1 - F_X(z)]^{n-1}$

❖ Distribution of the k-th maxima

- $f_{Y_k}(y) = \frac{n!}{(k-1)!(n-k)!} f_X(y)[F_x(y)]^{k-1}[1 - F_X(y)]^{n-k}$

▣ Order Statistics (continued)

❖ Extreme Value Distributions

➢ Distribution of smallest values

- Defining the random variable $\eta_n = nF_X(z)$, for u in $0 \le u \le n$
  - $\Gamma_n(u) = P(\eta_n \le u) = P(nF_X(z) \le u)$

$$= P\left(z \le F_X^{-1}\left(\frac{u}{n}\right)\right) \qquad \because F_X(z) \text{ is a monotonically increasing function}$$

$$= F_Z\left(F_X^{-1}\left(\frac{u}{n}\right)\right)$$

$$= 1 - \left[1 - F_X\left(F_X^{-1}\left(\frac{u}{n}\right)\right)\right]^n = 1 - \left(1 - \frac{u}{n}\right)^n$$

- As $n \to \infty$,
  - $\Gamma(u) = \lim_{n\to\infty} \Gamma_n(u) = 1 - e^{-u}, \quad u \ge 0$
  - $\gamma(u) = e^{-u}, \quad u \ge 0$
- Distribution of the minimum, z for a large n
  - Since η_n is a monotonically increasing function of z, $P(Z \le z) = P(\eta_n \le u)$
  - $F_Z(z) = \Gamma_n(u)$
  - For a large n, $F_Z(z) \cong 1 - e^{-u} = 1 - e^{-nF_X(z)}$

▣ Order Statistics

❖ Extreme Value Distributions

➢ Distribution of smallest values (continued)

- Example: X is a uniform deviate in $[0, A]$
  - $F_X(x) = x/A \;\rightarrow\; \eta_n = nF_X(z) = nz/A$
  - $F_Z(z) \cong 1 - e^{-u} = 1 - e^{-nz/A},\; z \geq 0$
  - $f_Z(z) \cong \frac{n}{A} e^{-nz/A},\; z \geq 0$

➢ Distribution of largest values

- Defining the random variable $\xi_n = n\big(1 - F_X(y)\big)$, for u in $0 \leq u \leq n$
  - $\Lambda_n(v) = P(\xi_n \leq v) = P\big(n\big(1 - F_X(y)\big) \leq v\big)$

$$= P(F_X(y) \geq 1 - v/n)$$

$$= P\left(y \geq F_X^{-1}\left(1 - \frac{v}{n}\right)\right) \quad \because F_X(y) \text{ is a monotonically increasing function}$$

$$= 1 - F_Y\left(F_X^{-1}\left(1 - \frac{v}{n}\right)\right)$$

$$= 1 - \left[F_X\left(F_X^{-1}\left(1 - \frac{v}{n}\right)\right)\right]^n = 1 - \left(1 - \frac{v}{n}\right)^n$$

▣ Order Statistics

❖ Extreme Value Distributions

➢ Distribution of largest values (continued)

- As $n \to \infty$,
  - $\Lambda(v) = \lim_{n\to\infty} \Lambda_n(v) = 1 - e^{-v}, \quad v \geq 0$
  - $\lambda(v) = e^{-v}, \quad v \geq 0$
- Distribution of the maximum, y for a large n
  - Since ξ_n is a monotonically decreasing function of y, $P(Y \leq y) = P(\xi_n \geq v)$
  - $F_Y(y) = 1 - \Lambda_n(v)$
  - For a large n, $F_Y(y) \cong 1 - \Lambda(v) = e^{-v} = e^{-n(1-F_X(y))}$
- Example: X is a exponential deviate in $[x_0, \infty]$
  - $F_X(x) = 1 - e^{-\beta(x-x_0)} \ \rightarrow \ \xi_n = n[1 - F_X(y)] = ne^{-\beta(y-x_0)}$
  - $F_Y(y) \cong e^{-v} = \exp\left[-ne^{-\beta(y-x_0)}\right], \ y \geq x_0$
  - $f_Y(y) \cong n\beta e^{-\beta(y-x_0)} \exp\left[-ne^{-\beta(y-x_0)}\right], \ y \geq x_0$

▣ Order Statistics

❖ Extreme Value Distributions (continued)

➢ Remark

- In the previous example for $F_X(x) = 1 - e^{-\beta(x-x_0)}$, we obtained the distribution of the largest value, $F_Y(y) \cong e^{-v} = \exp\left[-ne^{-\beta(y-x_0)}\right],\ y \geq x_0$
- We will derive the same results for the Poisson process
- The probability of $X \geq x$ is $p_x = 1 - F_X(x) = e^{-\beta(x-x_0)}$
- Assuming, during time t, the annual rate of events larger than x_0 is ν_0, the number of events larger than x is $\nu_0 p_x t$
- Then, the probability for the occurrence of n events with $X \geq x$ is

$$P[N = n] = \frac{(\nu_0 p_x t)^n e^{-\nu_0 p_x t}}{n!}$$

- The probability that the magnitude of the largest event, Y is less than x is equal to the probability that there is no event larger than x:
- $$\begin{aligned} F_Y(x) &= \exp[-\nu_0 p_x t] = \exp\left[-\nu_0 t e^{-\beta(x-x_0)}\right] \\ &= \exp\left[-ne^{-\beta(x-x_0)}\right] \quad \because n = \nu_0 t \end{aligned}$$
- Or $F_Y(y) = \exp\left[-ne^{-\beta(y-x_0)}\right],\ y \geq x_0$
- Here, we didn't assume a large number, but the Poisson process

♣ Generalized Extreme Value (GEV) Distribution

- Extreme value distribution (EVD) are classified into 3 types
  - Type I: Gumbel Distribution (also called the Gumbel-Type)
    - The most common EVD and has two forms: one for the minimum, and one for the maximum
    - It is defined in the unbounded range
  - Type II: Fréchet Distribution
    - Used to model maximum values in a data set
    - Its is bounded (restricted) on the lower side
  - Type III: Weibull Distribution
    - Used in assessing product reliability to model failure times and life data analysis
- GEV distribution unites all the 3 types of EVD above
  - $F(x;\mu,\sigma,\rho) = exp\left\{-\left[1+\rho\left(\frac{x-\mu}{\sigma}\right)\right]^{-1/\rho}\right\} = e^{-t(x)}$
  - An EVD type is determined by the (shape) parameter ρ
    - $\rho = 0$: Type I $\rightarrow t(x) = e^{-\frac{x-\mu}{\sigma}}$, $x \in (-\infty, +\infty)$
    - $\rho > 0$: Type II $\rightarrow t(x) = \left[1+\rho\left(\frac{x-\mu}{\sigma}\right)\right]^{-1/\rho}$, $x \in [\mu - \frac{\sigma}{\rho}, +\infty)$
    - $\rho < 0$: Type III $\rightarrow t(x) = \left[1+\rho\left(\frac{x-\mu}{\sigma}\right)\right]^{-1/\rho}$, $x \in (-\infty, \mu - \frac{\sigma}{\rho}]$

▣ One Function of Two Random Variables

❖ $Z = X + Y$

➢ $F_Z(z) = P(Z \le z)$

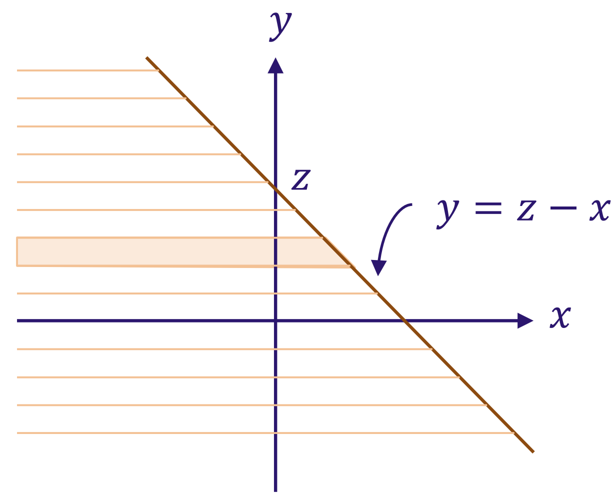
$$\cancel{= P(Z \le x + y)}$$

$$= P(X + Y \le z)$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{z-y} f_{X,Y}(x,y)dxdy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{z-y} f_X(x)f_Y(y)dxdy$$

$$= \int_{-\infty}^{\infty} F_X(z-y)f_Y(y)dy$$



➢ $f_Z(z) = \frac{\partial}{\partial z}\int_{-\infty}^{\infty} F_X(z-y)f_Y(y)dy$

$$= \int_{-\infty}^{\infty} f_X(z-y)f_Y(y)dy \leftrightarrow \int_{-\infty}^{\infty} f_{X,Y}(z-y,y)dy$$

❖ $Z = X/Y$

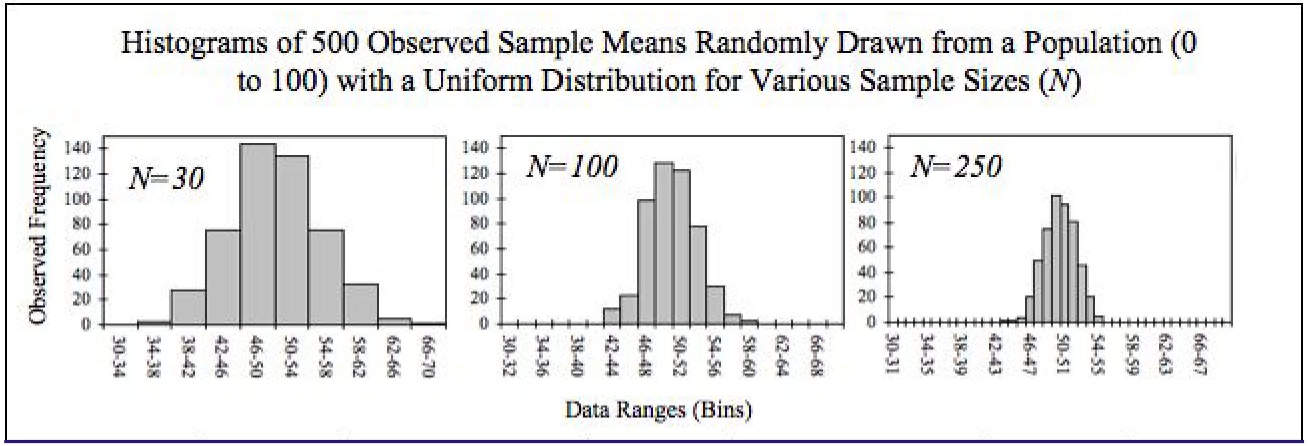
➢ $f_Z(z) = \int_{-\infty}^{\infty}|y|f_X(zy)f_Y(y)dy \leftrightarrow \int_{-\infty}^{\infty}|y|f_{X,y}(zy,y)dy$

▣ Central Limit Theorem

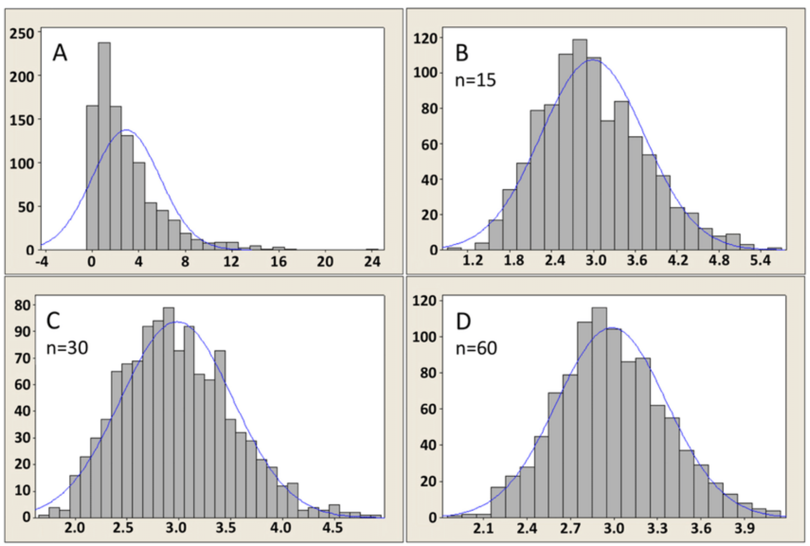
❖ Definition

- There are several versions of CLT
- In probability theory, CLT states that, under appropriate conditions, the distribution of a normalized version of the sample mean converges to a standard normal distribution. This holds even if the original variables themselves are not normally distributed.
- In statistics, CLT can be stated as: let $X_1, X_2, \cdots, X_n$ denote a statistical sample from a population with mean μ and variance σ^2, and let $\bar{X}_n$ denote the sample mean. Then as $n \to \infty$, the distribution of $\frac{(\bar{X}_n - \mu)}{\sigma/\sqrt{n}}$ is a normal distribution with mean 0 and variance 1.

❖ Explanation 1



<from fiveable>



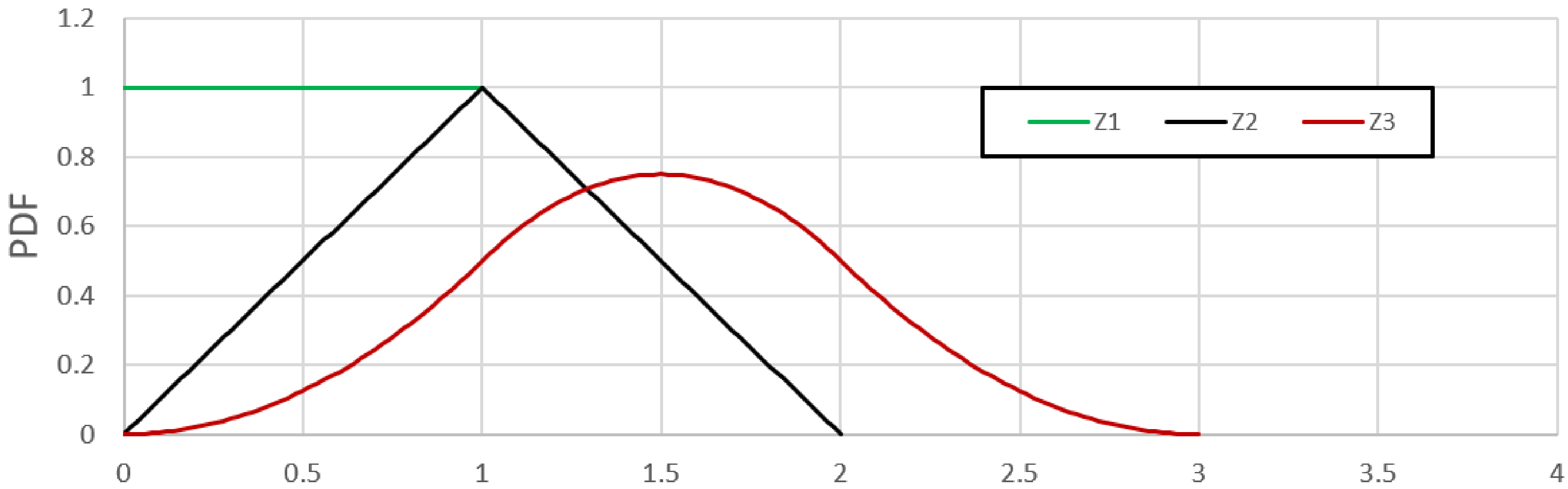
Panel A shows the population (highly skewed right and truncated at zero. Panel B, C, D show the distribution of sample means of sizes n=15, 30, and 60, respectively (from NCBI).

❖ Explanation 2

- For independent uniform deviates, $U_1, U_2, U_3, \cdots$ in [0,1]
  - $Z_1 = U_1, \quad f_{Z_1}(z) = \begin{cases} 1, & 0 \le z \le 1 \\ 0, & \text{otherwise} \end{cases}$
  - $Z_2 = U_1 + U_2, \quad f_{Z_2}(z) = \begin{cases} z, & 0 \le z \le 1 \\ 2 - z, & 1 \le z \le 2 \\ 0, & \text{otherwise} \end{cases}$
  - $Z_3 = U_1 + U_2 + U_3, \quad f_{Z_3}(z) = \begin{cases} \frac{1}{2} z^2, & 0 \le z \le 1 \\ -\left(z - \frac{3}{2}\right)^2 + \frac{3}{4}, & 1 \le z \le 2 \\ \frac{1}{2}(z-3)^2, & 2 \le z \le 3 \\ 0, & \text{otherwise} \end{cases}$
  - $Z_n = \sum_{i=1}^{n} U_i, \quad f_{Z_n}(z) \rightarrow \textit{normal distribution}$ as $n \rightarrow \infty$

❖ Explanation 2 (continued)

- Even Z_3 almost resembles a normal distribution
- Note that the range of z increases as increasing n



Chapter 2
Estimation

Introduction

▣ Sample

- ❖ Each sample X_i is a random variable
- ❖ Value x_i of a sample X_i is a realization of X_i
- ❖ The set $\{X_1, X_2, \cdots, X_n\}$ is called a random sample of X, of which size is n

▣ Point estimation

- ❖ The value of some parameter θ (i.e., mean or variance) can be estimated using a function of the random sample $\{X_1, X_2, \cdots, X_n\}$
- ❖ The function used to estimate θ, $\hat{\theta} = \hat{\theta}(X_1, X_2, \cdots, X_n)$ is called an estimator of θ, and said to be a point estimator
- ❖ If $E(\hat{\theta}) = \theta$, then $\hat{\theta}$ is called an unbiased estimator
- ❖ If the variance of $\hat{\theta}$ is smaller, then $\hat{\theta}$ is said to be more efficient
- ❖ If $\lim_{n\to\infty} P\{|\hat{\theta} - \theta| < \epsilon\} = 1$ for an arbitrary positive ϵ, then $\hat{\theta}$ is called an consistent estimator

▣ Interval estimation

❖ The point estimate may deviate from the actual parameter value

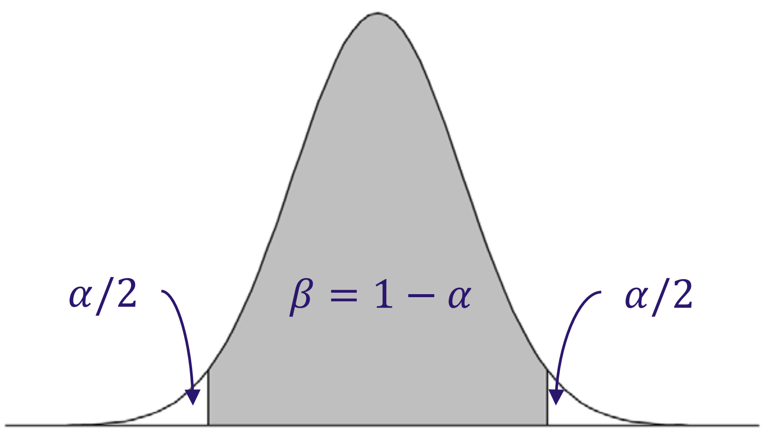
❖ To obtain an estimate with a high confidence, it is necessary to construct an interval estimate such that *the interval includes the actual parameter value with a high probability*

❖ Given an estimator $\hat{\theta}$, if

$$P\left(\hat{\theta} - e_1 < \theta < \hat{\theta} + e_2\right) = \beta$$

- ➢ The interval $(\hat{\theta} - e_1, \hat{\theta} + e_2)$ is said to be $100 \times \beta$ percent confidence interval for θ, and β is called the confidence coefficient or confidence level

❖ For a statistical test, it is more convenient to use $1 - \alpha$ in place of β



▣ Maximum Likelihood Method (MLM)

❖ Condition

- The functional form of the PDF of the random variable is known

❖ Assumption

- MLM is to choose an estimator such that the observed sample is the most likely to occur among all possible samples

❖ General properties

- Usually produces estimators that have minimum variance and consistency properties
- If the sample size is small, however, the estimator may be biased

❖ Formulation

- Assuming X has a PDF $f(x|\theta)$, where θ is an unknown parameter to be estimated,
- The likelihood function to be maximized over θ is

$$L(\theta) = \prod_{i=1}^{n} f(x_i|\theta)$$

▣ Maximum Likelihood Method (continued)

❖ Formulation (continued)

- Easier way is to work with log-likelihood

  $\ln L(\theta) = \sum_{i=1}^{n} \ln f(x_i|\theta)$

- Two conditions to maximize the likelihood function
  - $\frac{\partial}{\partial\theta} \ln L(\theta) = 0$ and $\frac{\partial^2}{\partial\theta^2} \ln L(\theta) < 0$
- Estimation of variance, for large n

  $Var(\hat{\theta}) = -\left[\frac{\partial^2}{\partial\theta^2} \ln L(\theta)\right]_{\theta=\hat{\theta}}^{-1}$

❖ Example

- Assuming X is exponentially distributed with a rate λ,
  - $L(\lambda) = \prod_{i=1}^{n} \lambda e^{-\lambda x_i} = \lambda^n \exp(-\lambda \sum_{i=1}^{n} x_i)$ or
  - $\ln L(\lambda) = n \ln \lambda - \lambda \sum_{i=1}^{n} x_i$
- Differentiating once and twice
  - $\frac{\partial}{\partial\lambda} \ln L(\lambda) = \frac{n}{\lambda} - \sum_{i=1}^{n} x_i$, $\frac{\partial^2}{\partial\lambda^2} \ln L(\lambda) = -\frac{n}{\lambda^2} < 0$

▣ Maximum Likelihood Method (continued)

❖ Example (continued)

➢ Setting the 1st derivative equal to 0, we have

- $\frac{\partial}{\partial\lambda}\ln L(\lambda) = \frac{n}{\lambda} - \sum_{i=1}^{n} x_i = 0$ or
- $\hat{\lambda} = \frac{n}{\sum_{i=1}^{n} x_i} = 1/\bar{x}$

➢ Using the 2nd derivative to calculate the variance of $\hat{\lambda}$

- $\frac{\partial^2}{\partial\lambda^2}\ln L(\lambda)|_{\lambda=\hat{\lambda}} = -\frac{n}{\hat{\lambda}^2} = -\frac{\bar{x}^2}{n}$
- $Var(\hat{\lambda}) = -\left[\frac{\partial^2}{\partial\lambda^2}\ln L(\lambda)\right]_{\lambda=\hat{\lambda}}^{-1} = \frac{n}{\bar{x}^2}$

▣ Method of Moments

- ❖ Advantages
  - ➢ The PDF needs not be in an explicit function of parameters
  - ➢ The procedure is fairly simple and the estimators are consistent
- ❖ Disadvantages
  - ➢ The estimators are often biased
- ❖ Definitions of moments
  - ➢ Population moments
    - $m_k = E(X^k) = \int x^k f_X(x|\theta)dx$
  - ➢ Sample moments
    - $\hat{m}_k = \frac{1}{n}\sum_{i=1}^{n}(x_i)^k$
  - ➢ Note that
    - the above definitions at centered at the origin
    - one can use the moments centered at the location (mean)

Method of Moments (continued)

Formulation

- If there are k parameters to be estimated, calculate the population moments and the sample moments up to the order k
- Second, solve the simultaneous equations

$$
\begin{aligned}
m_1 &= \hat{m}_1 \\
m_2 &= \hat{m}_2 \\
&\vdots \\
m_k &= \hat{m}_k
\end{aligned}
$$

Example

- If X is sampled from a gamma distribution, $X \sim G(b, c)$
  - $m_1 = \frac{b}{c}$; $\quad m_2 = \frac{b}{c^2} + \frac{b^2}{c^2}$
  - $\hat{m}_1 = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{X}$; $\quad \hat{m}_2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 = V^2 + (\bar{X})^2$
- Solving for b and c
  - $\hat{b} = \frac{(\bar{X})^2}{V^2}$; $\quad \hat{c} = \frac{\bar{X}}{V^2}$

▣ Least-Squares Method (LSM)

❖ Observation, prediction, and error

- The sample x_i can be regarded as the observation at z_i
- The model to predict observations is $g(X|\theta)$ where θ is a model parameter
- The error between the observation and the prediction is;

$$e_i = x_i - g(z_i|\theta)$$

❖ Sum of squared errors (SSE)

- $SSE = \sum_{i=1}^{n}(e_i)^2 = \sum_{i=1}^{n}(x_i - g(z_i|\theta))^2$
- The estimator $\hat{\theta}$ is the value of θ that minimizes the SSE, and obtained by solving;

$$\frac{\partial}{\partial\theta}SSE = 0 \text{ and } \frac{\partial^2}{\partial\theta^2}SSE > 0$$

❖ Example

- Prediction model: $g(z_i|\theta) = g(z_i|a,b) = az_i + b$
- Prediction error: $e_i = x_i - (az_i + b)$
- $SSE = \sum_{i=1}^{n}(x_i - az_i - b)^2$

▣ Least-Squares Method (continued)

❖ Example (continued)

➢ Parameters a, b that minimize the SSE are;

$$\frac{\partial SSE}{\partial a} = 0 \text{ and } \frac{\partial SSE}{\partial b} = 0;$$

$$\frac{\partial^2}{\partial a^2} SSE = 2\sum_{i=1}^{n} {z_i}^2 > 0 \text{ and } \frac{\partial^2}{\partial b^2} SSE = 2\sum_{i=1}^{n} 1^2 = n > 0$$

➢ Solving for a, b yields;

$$\hat{a} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(z_i - \bar{z})}{\sum_{i=1}^{n}(z_i - \bar{z})^2} \text{ and } \hat{b} = \bar{x} - \hat{a}\bar{z}$$

where $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ and $\bar{z} = \frac{1}{n}\sum_{i=1}^{n} z_i$

➢ Variances

- $Var(\hat{a}) = \frac{\sigma^2}{n} + \frac{\bar{z}^2\sigma^2}{\sum_{i=1}^{n}(z_i - \bar{z})^2} \quad \because \sigma^2 = Var(X) \approx \frac{\sum_{j=1}^{n}(x_j - \bar{x})^2}{n-1}$
- $Var(\hat{b}) = \frac{\sigma^2}{\sum_{i=1}^{n}(z_i - \bar{z})^2}$

▣ Least-Squares Method (continued)

❖ MLM equivalency

- If $e_i = x_i - g(z_i|\theta) \sim N(0, \sigma^2)$, the likelihood function for e_i is
- $L(\theta) = \left(\sqrt{2\pi}\sigma\right)^{-n} \prod_{i=1}^{n} e^{-\frac{{e_i}^2}{2\sigma^2}} = \left(\sqrt{2\pi}\sigma\right)^{-n} \exp\left(-\frac{SSE}{2\sigma^2}\right)$
- Maximization of $L(\theta)$ is equivalent to minimization of the exponent which is the least-squares
- Variances
  - $Var(\hat{a}) = \frac{\bar{z}^2\sigma^2}{\sum_{i=1}^{n}(z_i-\bar{z})^2}$ $\quad \because \sigma^2 = Var(X) \approx \frac{\sum_{j=1}^{n}\left(x_j-\bar{x}\right)^2}{n-1}$
  - $Var\left(\hat{b}\right) = \frac{\sigma^2}{\sum_{i=1}^{n}(z_i-\bar{z})^2}$

Least-Squares Method (continued)

❖ Weighted least-squares method (WLSM)

- If the errors are mutually independent, but not identically distributed, i.e., $e_i = x_i - g(z_i|\theta) \sim N(0, \sigma_i^2)$, then, from the MLM equivalency, the likelihood function for e_i becomes

  $$L(\theta) = \left(\sqrt{2\pi}\sigma_i\right)^{-n} \exp\left(-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{e_i}{\sigma_i}\right)^2\right) = \left(\sqrt{2\pi}\sigma_i\right)^{-n} \exp\left(-\frac{1}{2}\sum_{i=1}^{n}(w_i e_i)^2\right)$$

  where $w_i = 1/\sigma_i$, the weight of the i-th error
- The above equation states that the observation with larger variance, i.e., more uncertain observation, is less weighted
- Maximization of $L(\theta)$ can be achieved by minimizing $\sum_{i=1}^{n}\left(\frac{e_i}{\sigma_i}\right)^2$
- χ^2 estimation
  - Since $e_i \sim N(0, \sigma_i^2)$, $X = \sum_{i=1}^{n}\left(\frac{e_i}{\sigma_i}\right)^2 \sim \chi^2(n-m)$, where m is the number parameter in θ
  - This can be used to test the suitability of the model, the assumption of normality, or the data credibility (rule of thumb: $X \cong n - m$; mean of χ^2)

Chapter 3
Hypothesis Test

Introduction

▣ Statistical Hypotheses

❖ Null hypothesis, H_0

- A statistical hypothesis that is to be tested
- No significance difference between the populations specified in the experiments

❖ Alternative hypothesis, H_1

- Alternative to the null hypothesis
- There exits sufficient evidence to support the credibility of the alternative hypothesis

▣ Error Types

Table of error types		Null hypothesis, H_0	
		True	False
Decision about null hypothesis, H_0	Not reject	Correct inference	Type II error
	Reject	Type I error	Correct inference

Introduction

▣ Statistical Hypotheses

❖ Null hypothesis, H_0

- No significance difference between the populations specified in the experiments

❖ Alternative hypothesis, H_1

- There exits sufficient evidence to support the credibility of the alternative hypothesis

❖ Important!

- H_0 & H_1 should satisfy the MECE principle
- Test is carried out for H_0

▣ Error Types

Table of error types		Null hypothesis, H_0	
		True	False
Decision about null hypothesis, H_0	Not reject	Correct inference	Type II error
	Reject	Type I error	Correct inference

▣ Test Procedure

❖ Minimization of Errors

- Impossible to minimize both of type I and type II errors at the same time
- The statistical decision is based on the minimization of the type I error

❖ Significance Level, α

- Maximum allowed probability to commit the type I error

❖ Test statistic

- A quantity derived from the sample for statistical hypothesis testing
  - Ex: sample mean, sample variance

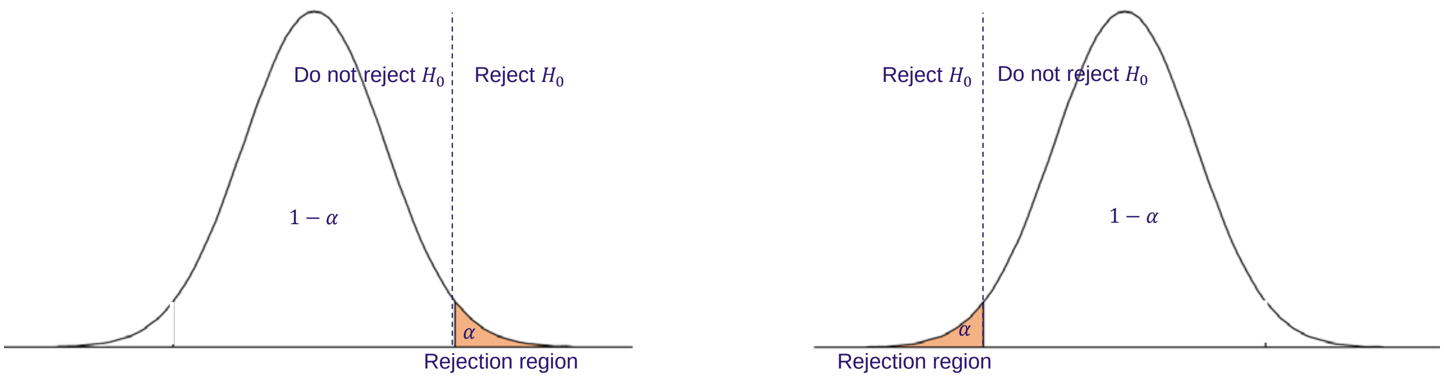
❖ Rejection region (critical region)

- a set of values for the test statistic for which the null hypothesis is rejected
- i.e., if the observed test statistic is in the critical region then we reject the null hypothesis and accept the alternative hypothesis
- It is determined per the alternative hypothesis

▣ Test Procedure (continued)

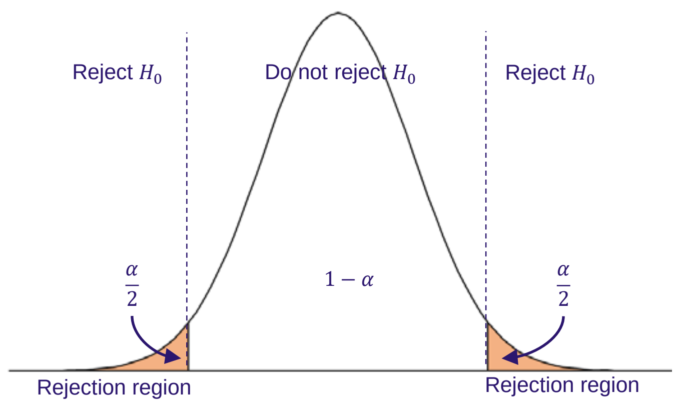
❖ One-sided test

- ➢ All the rejection region corresponding to the significance level is located at the lower end or upper end of the distribution



❖ Two-sided test

- ➢ The rejection region locates at two ends by half and half



▣ Test Procedure (continued)

❖ p-value

- The probability that the test statistic is exceeded or falling short
- The one-ended test
  - When the rejection region is at the upper tail
    - p-value is the probability to exceed the statistic
    - the null hypothesis is rejected if p-value is smaller than the significance level α
  - When the rejection region is at the lower tail
    - p-value is the probability not to exceed the statistic
    - the null hypothesis is rejected if p-value is smaller than the significance level α
- The two-ended test
  - If p-value is greater than the significance level $\alpha/2$ or smaller than $1-\alpha/2$, the null hypothesis is rejected

▣ Test of Population Mean Estimate

❖ Distribution of sample mean

- For $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$, we know that $E(\bar{X}) = \mu$ and $Var(\bar{X}) = \frac{\sigma^2}{n}$
- For large n, from CLT, $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$, or $Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$ if the variance of population σ^2 is known
- If σ^2 is unknown, using the sample variance,
  - If n is reasonably large (i.e., larger than 30), then $\frac{\bar{X}-\mu}{\sqrt{V/n}} \sim N(0,1)$
  - If n is smaller than 30
    - If X_i is a normal deviate, then $(\bar{X}-\mu)/(\sigma/\sqrt{n}) \sim N(0,1)$ and $\frac{(n-1)V}{\sigma^2} \sim \chi^2(n-1)$ so that

      $$\frac{\bar{X}-\mu}{\sqrt{V/n}} = \frac{(\bar{X}-\mu)/(\sigma/\sqrt{n})}{\sqrt{\frac{(n-1)V}{\sigma^2}/(n-1)}} \sim t(n-1)$$
    - If X_i is an exponential deviate, then $2n\bar{X}/\mu \sim \chi^2(2n)$

Test Examples

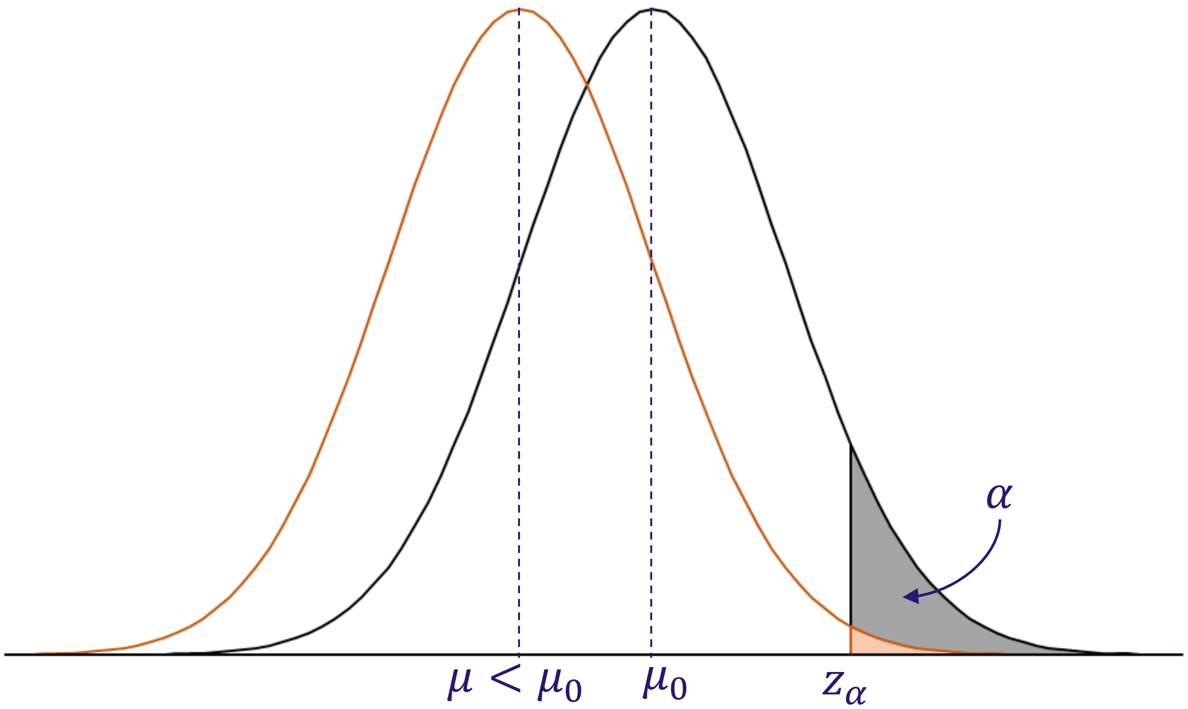
❖ Test statistic: $T = \frac{\bar{X}-\mu_0}{\sqrt{V/n}}$

- For significance level α

Null Hypothesis	Rejection Region		
	$n \geq 30$	$n < 30$	
		Normal X_i	Exponential X_i
$H_0: \mu \leq \mu_0$ $H_1: \mu > \mu_0$	$\frac{\bar{X}-\mu_0}{\sqrt{V/n}} > z_\alpha$	$\frac{\bar{X}-\mu_0}{\sqrt{V/n}} > t_\alpha(n-1)$	$\frac{2n\bar{X}}{\mu_0} > \chi^2_\alpha(2n)$
$H_0: \mu \geq \mu_0$ $H_1: \mu < \mu_0$	$\frac{\bar{X}-\mu_0}{\sqrt{V/n}} < -z_\alpha$	$\frac{\bar{X}-\mu_0}{\sqrt{V/n}} < -t_\alpha(n-1)$	$\frac{2n\bar{X}}{\mu_0} < \chi^2_{1-\alpha}(2n)$
$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$	$\left\lvert\frac{\bar{X}-\mu_0}{\sqrt{V/n}}\right\rvert > z_{\alpha/2}$	$\left\lvert\frac{\bar{X}-\mu_0}{\sqrt{V/n}}\right\rvert > t_{\alpha/2}(n-1)$	$\frac{2n\bar{X}}{\mu_0} > \chi^2_{\alpha/2}(2n)$ or $\frac{2n\bar{X}}{\mu_0} < \chi^2_{1-\alpha/2}(2n)$

z_α: a value of the standard normal deviate of which probability to exceed it is α

V: sample variance



α
$\mu < \mu_0$
μ_0
z_α

Distribution of Sample Mean & Variance

▣ Sample Variance

❖ For unknown population mean μ, we have $V = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$

❖ If $\frac{X_i-\mu}{\sigma} \sim N(0,1)$, then $\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2 \sim \chi^2(n)$

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

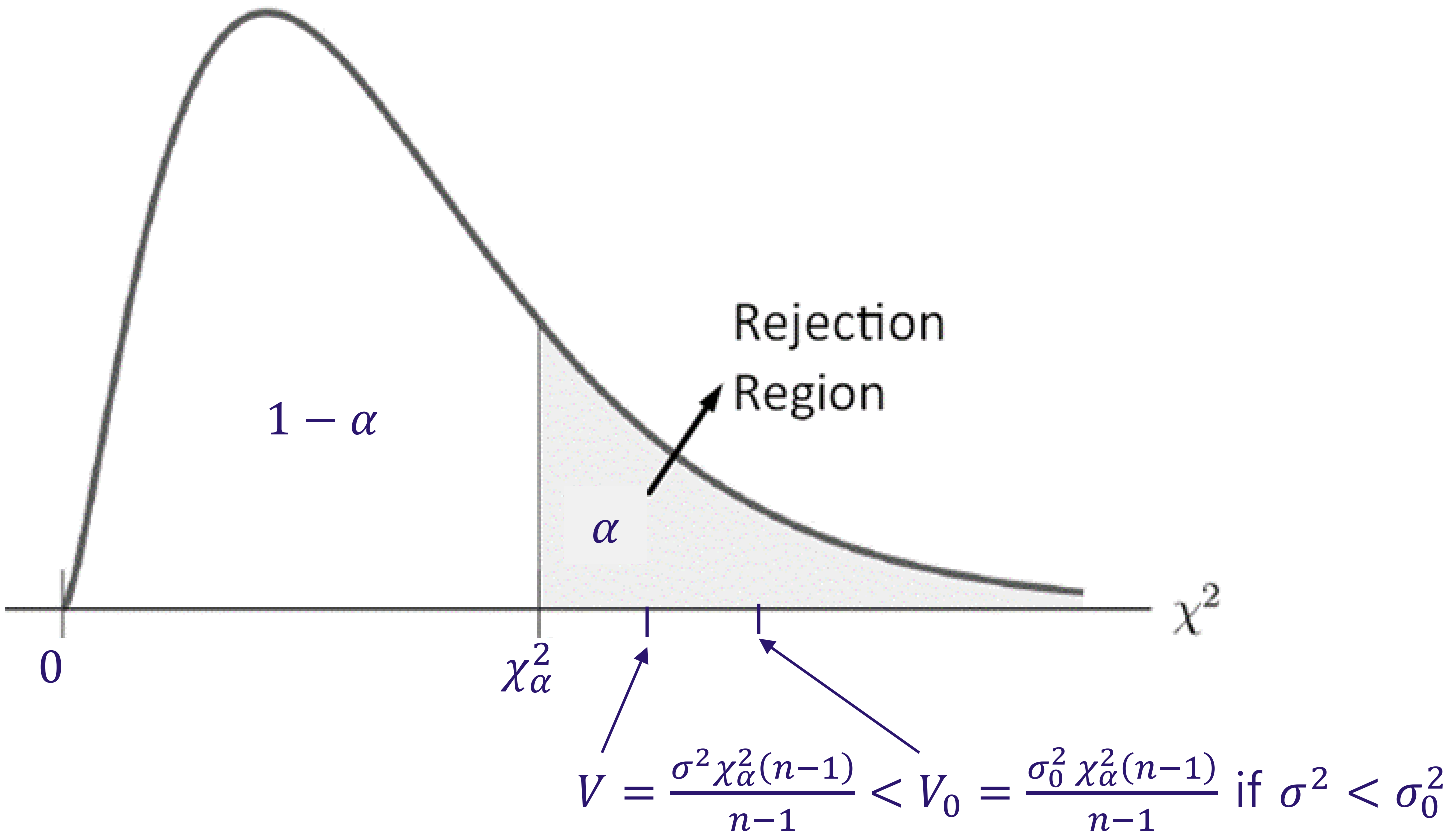
$$= \frac{(n-1)V}{\sigma^2} + \left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right)^2$$

❖ Since $\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right)^2 \sim \chi^2(1)$,it follows that $\frac{(n-1)V}{\sigma^2} \sim \chi^2(n-1)$

- Test statistic: $T = \frac{(n-1)V}{\sigma_0^2}$
  - For significance level α
- If X_i follows the normal distribution

Null Hypothesis	Rejection Region
$H_0: \sigma^2 \le \sigma_0^2$	$\frac{(n-1)V}{\sigma_0^2} > \chi_\alpha^2(n-1)$
$H_0: \sigma^2 \ge \sigma_0^2$	$\frac{(n-1)V}{\sigma_0^2} < \chi_{1-\alpha}^2(n-1)$
$H_0: \sigma^2 \ne \sigma_0^2$	$\frac{(n-1)V}{\sigma_0^2} > \chi_{\alpha/2}^2(n-1)$ or $\frac{(n-1)V}{\sigma_0^2} < \chi_{1-\alpha/2}^2(n-1)$

$\chi_\alpha^2(n-1)$: a value of the Chi-square deviate of $(n-1)$ degrees of freedom, of which probability to exceed it is α


Rejection
Region
1 − α
α
0
χ^2_α
χ^2
$V = \frac{\sigma^2\chi^2_\alpha(n-1)}{n-1} < V_0 = \frac{\sigma_0^2\,\chi^2_\alpha(n-1)}{n-1}$ if $\sigma^2 < \sigma_0^2$

▣ Test of Distributions

❖ Chi-square test

- ➢ Used for the grouped data
- ➢ Pearson's test statistic: $PTS = \sum_n^N \frac{(O_n - E_n)^2}{E_n} \sim \chi^2(N-M)$
  - O_n : observed frequency
  - E_n : expected frequency from the assumed distribution
  - M = 1+constraints related to estimation of parameters of the distribution

❖ Kolmogorov-Smirnov test

- ➢ Used for the continuous data
- ➢ Test statistic: $D = max|S(x_n) - F(x_n)|,\ \ n = 1, 2, \cdots, N$
  - x_n : observation in ascending order
  - $S(x_n) = n/N$: empirical cumulative distribution
  - $F(x_n)$: cumulative distribution of the assumed distribution
  - $\Pr(D > d) = Q(\sqrt{N}d)$
    - $Q(x) = 2\sum_{j=1}^{\infty}(-1)^{j-1}e^{-2j^2x^2}$

❖ Shapiro-Wilk test: specific to the test of the normality

▣ Examples

❖ Average (mean) lifetime of bulbs

- Situation: a company states that the average lifetime of their bulbs is longer than 1950 h
- Task: given the $n = 9$ samples with $\bar{X} = 1966.7$ and $V = 69.6^2$, test the hypothesis with the significance level 0.05

① Test statistic

- $T = \frac{\bar{X}-\mu}{\sqrt{V/N}}$

② Distribution of test statistic for $n = 9\ (< 30)$:

- $T = \frac{\bar{X}-\mu}{\sqrt{V/N}} \sim t(n-1) = t(8)$

③ Hypotheses

- H_0: $\mu \le \mu_0 = 1950$ $\qquad H_1$: $\mu > \mu_0 = 1950$

④ Rejection region

- $\tau = \frac{\bar{X}-\mu_0}{\sqrt{V/N}} = \frac{\bar{X}-1950}{\sqrt{69.6/9}} = 0.720$
- Since $t_{0.05}(8) = 1.86 > \tau = 0.720$, H_0 cannot be rejected.

▣ Examples (continued)

❖ Variance of thickness of window glasses

- Situation: an investigator reports $\sigma^2 > 1.5^2$ due to malfunction of machines
- Given the n =10 samples with the sample variance $v = 5.1556$ and the thickness follows the normal distribution, test the report with the significance level 0.05

① Test statistic

- $X = \frac{(n-1)V}{\sigma^2}$

② Distribution of test statistic for n =10,

- $\frac{(n-1)V}{\sigma^2} \sim \chi\,(n-1) \rightarrow \frac{9V}{\sigma^2} \sim \chi\,(9)$

③ Hypotheses

- $H_0\colon \sigma^2 \leq 1.5^2\ (= \sigma_0^2)$ $\qquad H_1\colon \sigma^2 > 1.5^2$

④ Rejection region

- $x = \frac{(n-1)v}{\sigma_0^2} = \frac{9\times 5.1556}{1.5^2} = 20.6224$ and $\chi^2_{0.05}(9) = 16.919$
- Since $x^2 > \chi^2_{0.05}(9)$, H_0 is rejected.

▣ Examples (continued)

❖ Poisson process of earthquakes (Noh, 2016)

➢ By earthquake frequency

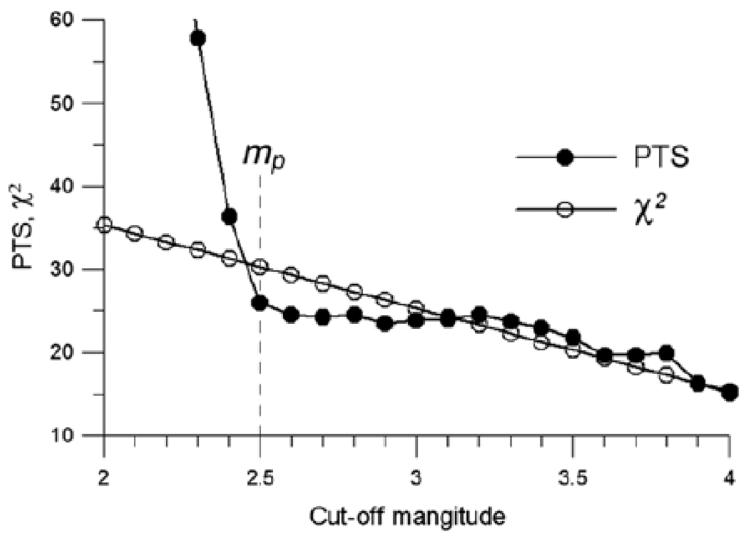
- H_0: earthquake frequency follows the Poisson process
- $\Pr(N = n) = \frac{(\lambda t)^n e^{-\lambda t}}{n!}$
  - t : exposure time; λ : mean annual rate
- Test statistic: $PTS = \sum_{n=0}^{N} \frac{(O_n - E_n)^2}{E_n} \sim \chi^2(N-2)$
  - O_n : observed frequency of time intervals in which earthquakes occurred n times
  - E_n : expectation of O_n, i.e., $E_n = \Pr(N = n) \times (\# \, of \; time \; intervals)$
  - $M = 2$: 1 + a constraint related to estimation of λ
- H_0 is rejected if $PTS > \chi^2_\alpha(N-1)$

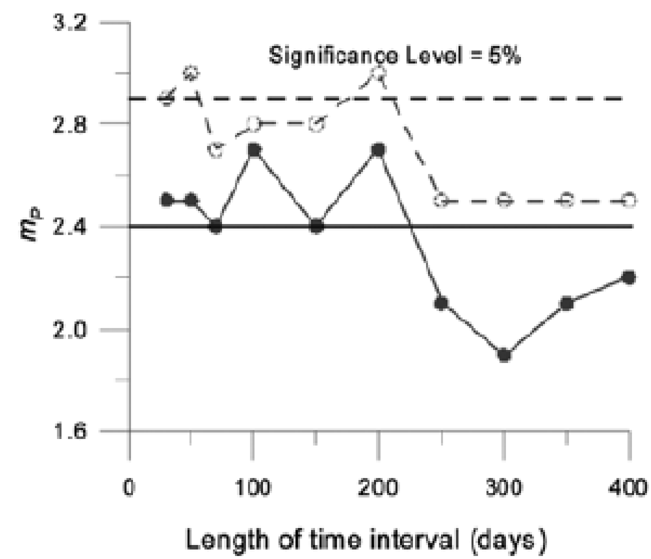
➢ By inter-event time

- $\Pr\big(N = 0\big) = e^{-\lambda t} = \Pr\big(T > t\big) = 1 - F(t)$
- Test statistic: $D = max|S(t_i) - F(t_i)|, \;\; i = 1, 2, \cdots, n$
  - t_i : observed inter-event time in ascending order
  - $S(t_i) = i/n$: empirical cumulative distribution
- $\Pr(D > d_{obs}) = Q(\sqrt{n}\, d_{obs})$
  - $Q(\epsilon) = 2\sum_{j=1}^{\infty}(-1)^{j-1} e^{2j^2\epsilon^2}$
  - H_0 is rejected if $Q(\sqrt{n}\, d_{obs}) < \alpha$

▣ Examples

❖ Poisson process of earthquakes (Noh, 2016))





Chapter 4
Monte Carlo Simulation

What is the Monte Carlo Simulation?

▣ Definition 1

- ❖ A statistical technique used to model and analyze the impact of uncertainty and variability in complex systems or processes
- ❖ It involves running a large number of simulations to estimate possible outcomes and their probabilities, often when the problem involves randomness or uncertainty

▣ Definition 2

- ❖ A computational technique used to model and analyze systems or processes that involve uncertainty, randomness, or complex variables
- ❖ It leverages random sampling and statistical analysis to approximate numerical results, often for problems that are difficult or impossible to solve analytically

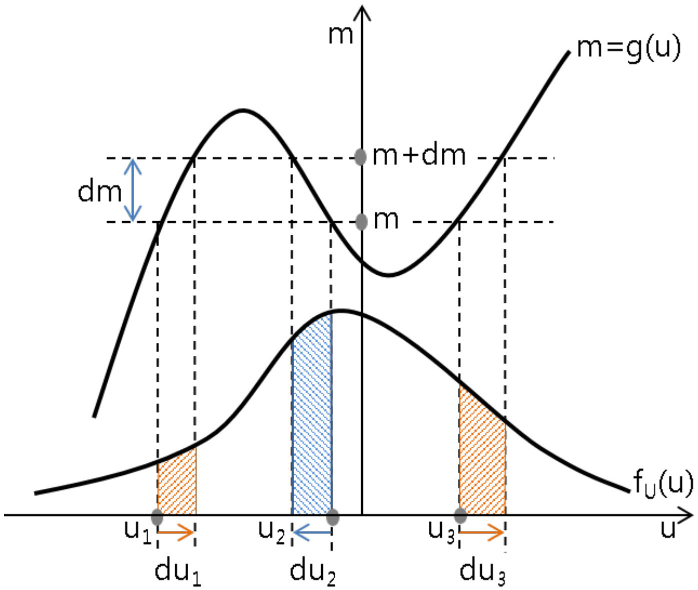
▣ Parametric Function

❖ $M = g(U)$

- where M and U are random variable

❖ Transform from $f_U(u)$ to $f_M(m)$

- $P(m < M < m + dm) = P(u_1 < U < u_1 + du_1) + P(u_2 + du_2 < U < u_2) + P(u_3 < U < u_3 + du_3)$ $\because du_1, du_3 > 0; du_2 < 0$
  - $P(m < M < m + dm) = f_M(m)dm$
  - $P(u_i < U < u_i + du_i) = f_U(u_i)|du_i|,$
  - $f_M(m)dm = f_U(u_1)|du_1| + f_U(u_2)|du_2| + f_U(u_3)|du_3|$
- $f_M(m) = \frac{f_U(u_1)}{|g\prime(u_1)|} + \frac{f_U(u_2)}{|g\prime(u_2)|} + \frac{f_U(u_3)}{|g\prime(u_3)|}$



▣ Parametric Function (continued)

❖ Example: $M = e^U \rightleftharpoons u = \ln m$

- One-to-one correspondence → $|g'(u)| = e^u = m > 0$
- $f_M(m) = \begin{cases} 0, & m \le 0 \\ \frac{1}{m} f_U(\ln m), & m > 0 \end{cases}$
- If $U \sim N(\mu, \sigma^2)$, i.e., $f_U(u) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(u-\eta)^2}{2\sigma^2}}$
  - $f_M(m) = \frac{1}{m\sqrt{2\pi}\sigma} e^{-\frac{(\ln m - \eta)^2}{2\sigma^2}}$
    - $\eta_M = e^{\eta + \sigma^2/2}$
    - $\sigma_M^2 = (e^{\sigma^2} - 1)e^{2\eta + \sigma^2}$

Monte Carlo Simulation

▣ Transform Method

❖ Uniform deviates in [0,1]

- $f_U(u) = 1,\ F_U(u) = u,\ \text{for}\ 0 \le u \le 1$

❖ Parametric function: $u = F_M(m)$

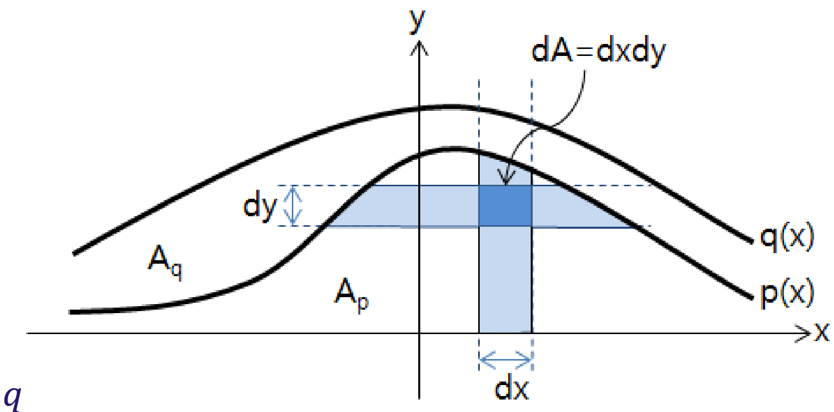
- Monotonically increasing function → solution at a single point
  - $0 \le u = F_M(m) \le 1$
  - $f_U(u) = \frac{f_M(m)}{|F_M'(m)|} = 1$

  ♣ U is an uniform deviate
- $M = F_M^{-1}(U)$

❖ Example

- $f_M(m) = k\beta e^{-\beta(m-m_{min})},\ \ m_{min} \le m \le m_{max}$
- $F_M(m) = k\left[1 - e^{-\beta(m-m_{min})}\right],\ \ m_{min} \le m \le m_{max}$
- $M = F_M^{-1}(U) = m_{min} - \ln\left(1 - \frac{U}{k}\right)/\beta,\ 0 \le U \le 1$

▣ Rejection Method



- ❖ Target function: $p(x)$
  - ➢ Target PDF: $f_p(x) = p(x)/A_p$
- ❖ Comparison function: $q(x)$
  - ➢ $q(x) \geq p(x), \quad \forall x$
  - ➢ Comparison PDF: $f_q(x) = q(x)/A_q$
  - ➢ $F_q^{-1}(u)$ is an explicit function
- ❖ Goal
  - ➢ To generate pairs of i.i.d. random variables (X, Y) that uniformly distribute between $q(x)$ and x-axis
- ❖ For independent uniform deviates U_1, U_2
  - ➢ $x = F_q^{-1}(u_1) \rightarrow P(x \leq X \leq x + dx) = f_q(x)dx = \frac{q(x)}{A_q}dx$ (1)
  - ➢ $y = q(x)u_2 \rightarrow P(y \leq Y \leq y + dy | x \leq X \leq x + dx) = \frac{dy}{q(x)}$ (2)
    - ▪ y is a uniform deviate in $[0, q(x)] \rightarrow f_Y(y) = \frac{1}{q(x)}$; constant, given an x
  - ➢ $P(dY, dX) = P(dY|dX)P(dX) = \frac{dy}{q(x)} \cdot \frac{q(x)}{A_q}dx = \frac{dxdy}{A_q}$

▣ Rejection Method (continued)

❖ Simulation procedure

① Generate a uniform deviate u_1 to get x by (1)

② Generate a uniform deviate u_2 to get y by (2)

③ Take x if $y \leq p(x)$, otherwise discard x

④ Repeat to get the necessary amount of x's

❖ Useful tip

- For the sake of efficiency, take the smallest A_q as far as it meets the condition $q(x) = f_q(x)A_q \geq p(x)$ at all x

▣ Examples

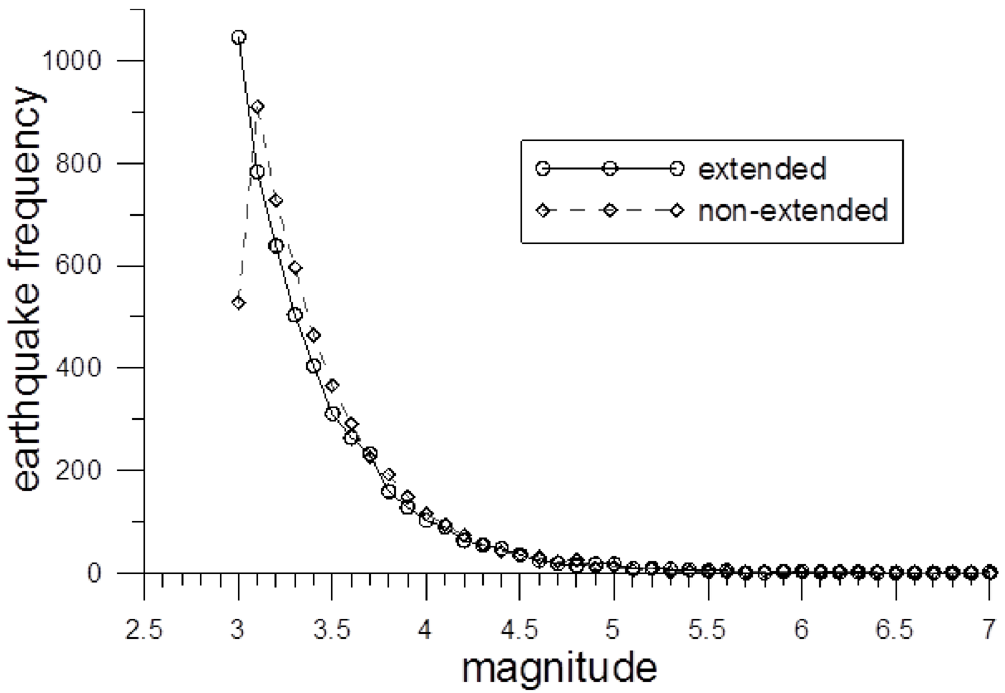
❖ Transform method for a complete catalog (Noh, 2014)

➢ $M = F_M^{-1}(U) = M_{min} - \ln\left(1 - \frac{U}{k}\right)/\beta,\ 0 \le U \le 1$

➢ Magnitude grouping

- $[m_{min}, m_{max}] \rightarrow [m_{min} - \Delta m/2\,, m_{max} + \Delta m/2]$

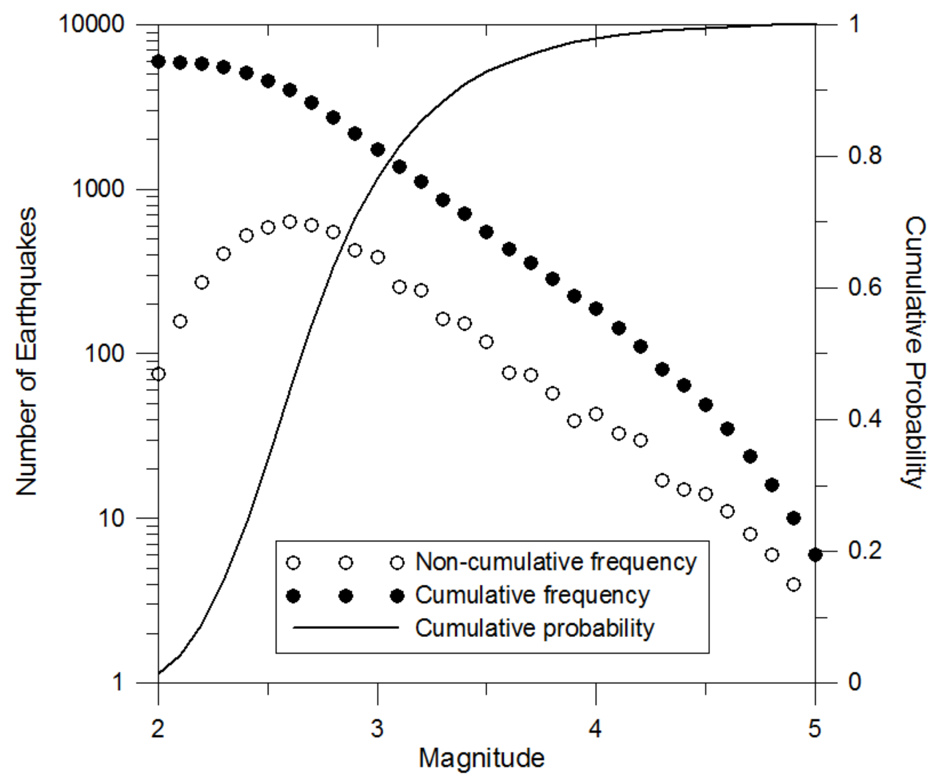
- $m_{min} = 3.0$
- $m_{max} = 7.0$
- $b = 1.0$
- $n_e = 5{,}000$



Examples (continued)

❖ Rejection method for a incomplete catalog (Noh, 2019)

- Target function: $p(m) = d(m) f_M(m)$
  - Detection rate: $d(m) = \begin{cases} c(m_c) \cdot \mathrm{erf}(m|\mu,\sigma) & \text{for } m < m_c \\ 1, & \text{for } m_c \le m \end{cases}$
    - $c(m_c) = 1/\mathrm{erf}(m_c|\mu,\sigma)$
- Comparison function
  - $q(m) = f_M(m) \ge p(m)$
  - $m_{min} = 2.0,\ m_{max} = 5.0$
  - $b = 1.0$
  - $m_c = 3.0$
  - $\mu = 5.0,\ \sigma = 0.25$
  - $n_e = 5{,}000$



지진원 특성평가

Characterization of Seismic Sources

- Part II -

Chapter 5
Earthquake Catalog

Preparation of Catalog

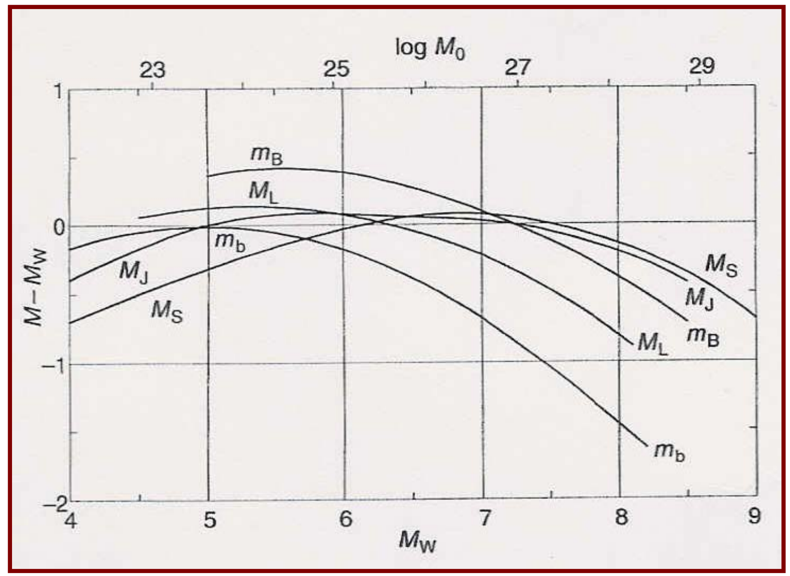
▣ Origin Parameters

- (Origin) Time
  - Time of earthquake occurrence
  - usually corresponds to the rupture initiation time
- Location
  - Locus at which an earthquake occurred, hypocenter
    - Usually corresponds to the rupture initiation point
  - Also described by epicenter and depth
  - Epicenter
    - Vertical projection of hypocenter to the surface
    - Described in geographical latitude & longitude
  - Focal depth
    - Depth (km) to the hypocenter
  - Distances to an earthquake
    - Hypocentral distance (d_H), epicentral distance (d_E), focal depth (h)
    - $d_H^2 = d_E^2 + h^2$; not a propagation distance

▣ Origin Parameters (continued)

❖ Size

- Various magnitude scales being used
- Body-wave magnitude
  - Sensitive to high-frequency content → larger value for deeper event
  - Saturated for large earthquakes
- Surface-wave magnitude
  - Measure of longer period energy → smaller value for smaller event
- Moment magnitude
  - Measure zero-frequency energy
  - No saturation
  - Physics-based value
  - Representative measure of size
  - $M = \frac{2}{3} log M_0 - 10.7$
    - M_0: seismic moment in dyne-cm
  - Bridge connecting to geology
    - $M_0 = \mu A \bar{D}$



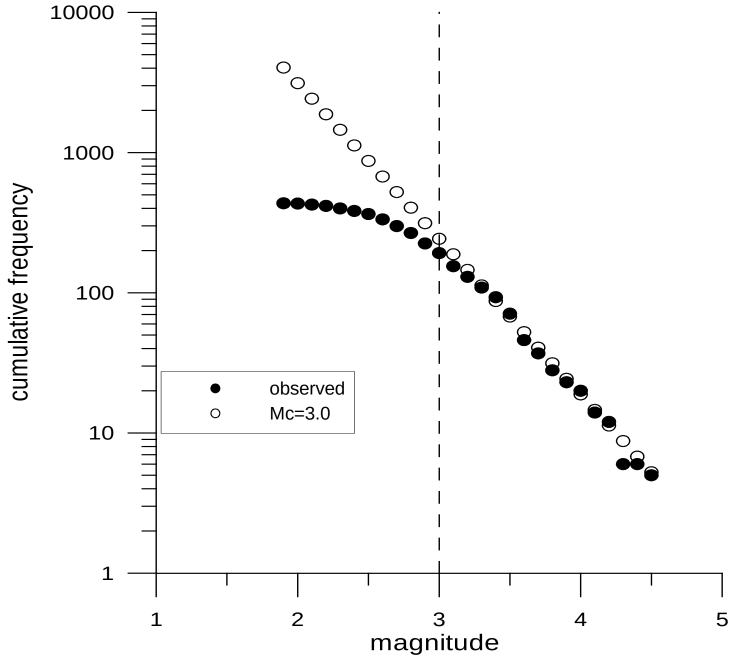
Integration of Catalogs

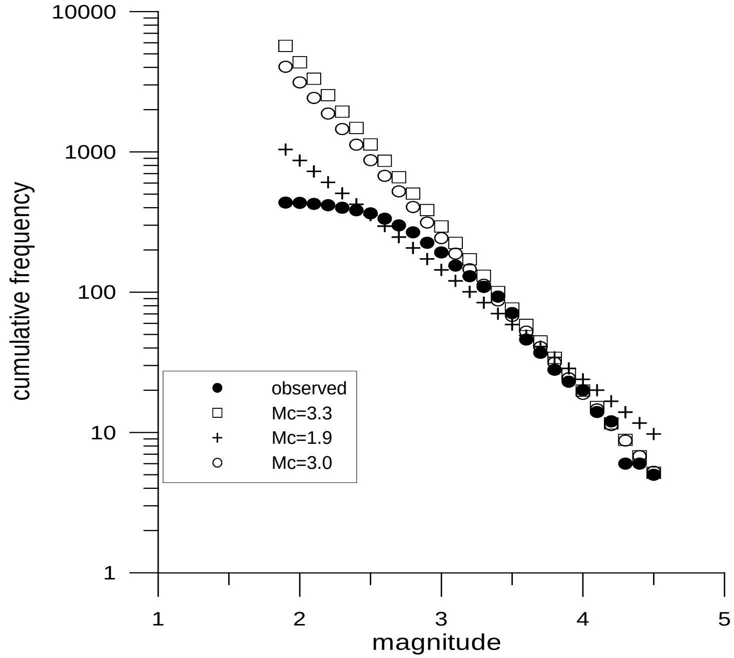
- To expand earthquake data by combining various catalogs covering different periods, different regions (ergodic assumption)
- General requirements
  - Description by unified quantities
  - Accuracy assessment
  - Completeness assessment
- Important properties to be checked
  - Unification of description
    - origin times
      - Use of UTC (Coordinated Universal Time) or a single local time
    - Unification of magnitude scale
      - Use of a single magnitude scale: moment magnitude is preferred
  - Accuracy checks
    - Error range of magnitude
    - Error range of location
  - Completeness checks
    - Completeness magnitudes of integrated catalogs

Completeness Assessment of Catalog

▣ Background

❖ The very 1st step of any analysis using earthquake catalog is to assess the completeness of the catalog at hand





➢ b=0.78, 1.11, 1.17 for m_c=1.9, 3.0, 3.3 [노명현 외(2000)]

Categories of Assessment Methods

- Network-based methods
  - Use detection capability of a seismic network
    - Background noise, network configuration, instrumentation, etc.
- Catalog-based methods
  - Use day-to-night noise modulation
    - Proposed by Rydelek and Sacks (1989)
    - Can be considered as a network-based method
  - Assumption of self-similarity for earthquake frequencies
    - $\log N = a - bM$
    - Focus of this course

General Procedure

- Introducing the cut-off magnitude, m_{co}
  - Starting from minimum magnitude of catalog
  - Gradually increasing by magnitude interval width
- Repetition of analysis for increasing m_{co}
- $m_c = m_{co}$ if certain conditions are met

▣ General Procedure (continued)

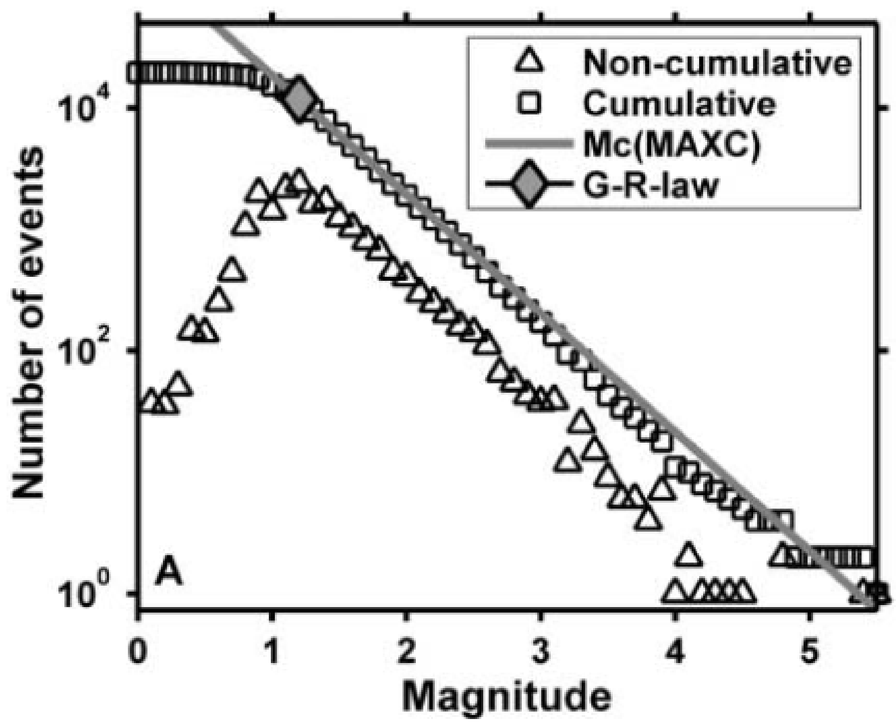
- ❖ Repetition of the above procedure to estimate m_c for bootstrap replicas of the catalog
- ❖ Calculation of the location and scale of m_c for the replicas

♣ Definitions of m_c

- ❖ Minimum magnitude above which all earthquakes were completely reported (Rydelek and Sacks, 2000)
- ❖ Minimum magnitude that preserves the information on seismicity parameters, i.e., m_{max}, Richter-b (Noh, 2019)

▣ Maximum Curvature Method

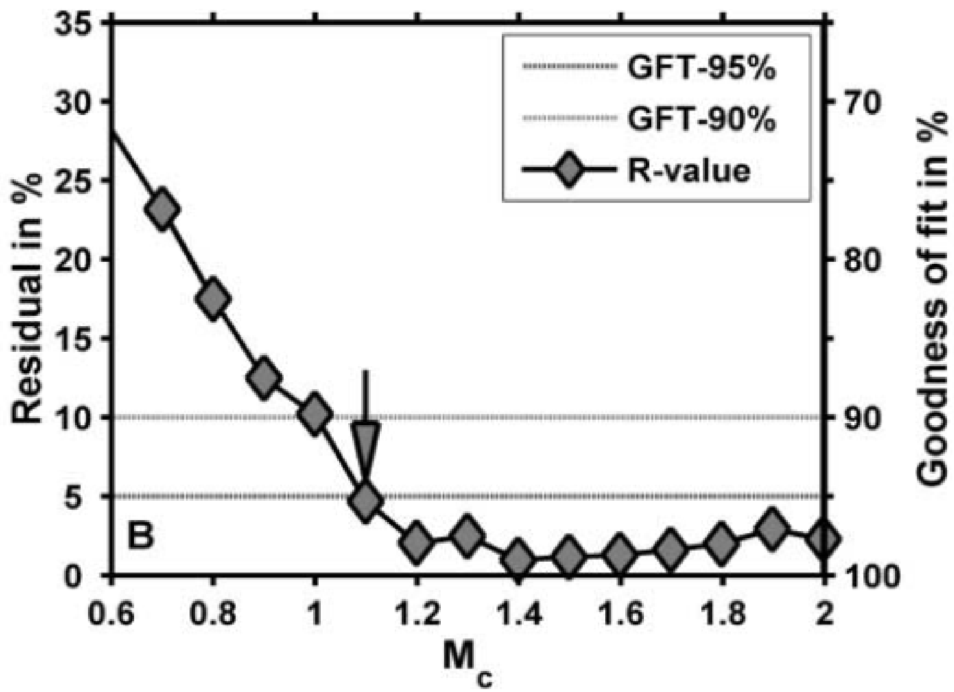
- ❖ Wiemer and Wyss (2000)
- ❖ m_c is located at the maximum of the non-cumulative frequency
- ❖ Simplest method, underestimation of m_c by 0.2



▣ Goodness-of-Fit Test

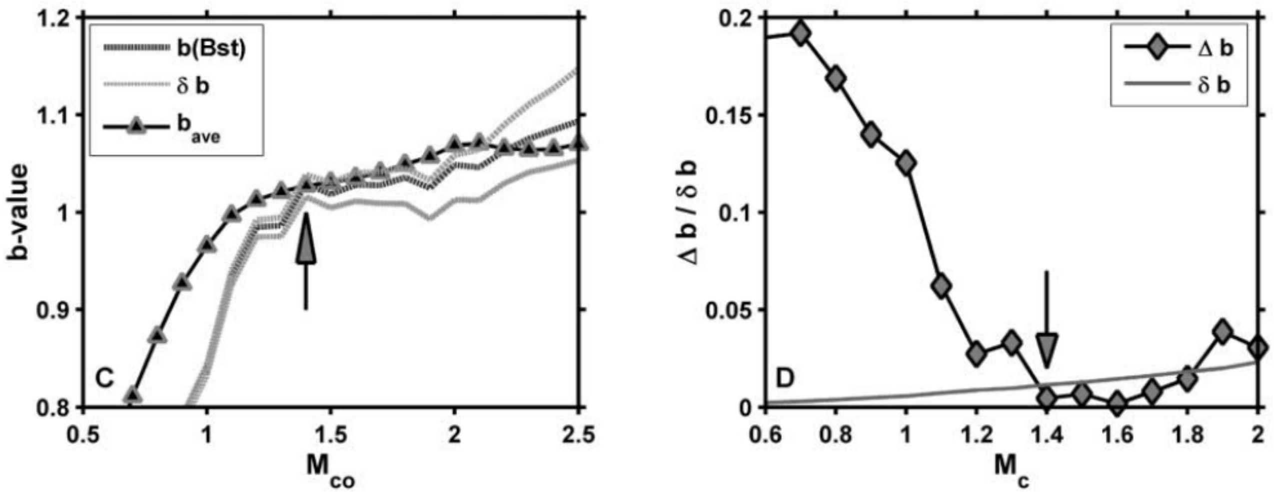
- ❖ Wiemer and Wyss (2000)
- ❖ $GFT(a,b,M_i) = 100 - \left(\frac{\sum_{M_i}^{M_{max}} \left| N_i^{obs} - N_i^{pre} \right|}{\sum N_i^{obs}} \times 100 \right)$

  where $\log N_i^{pre} = a - bM_i$



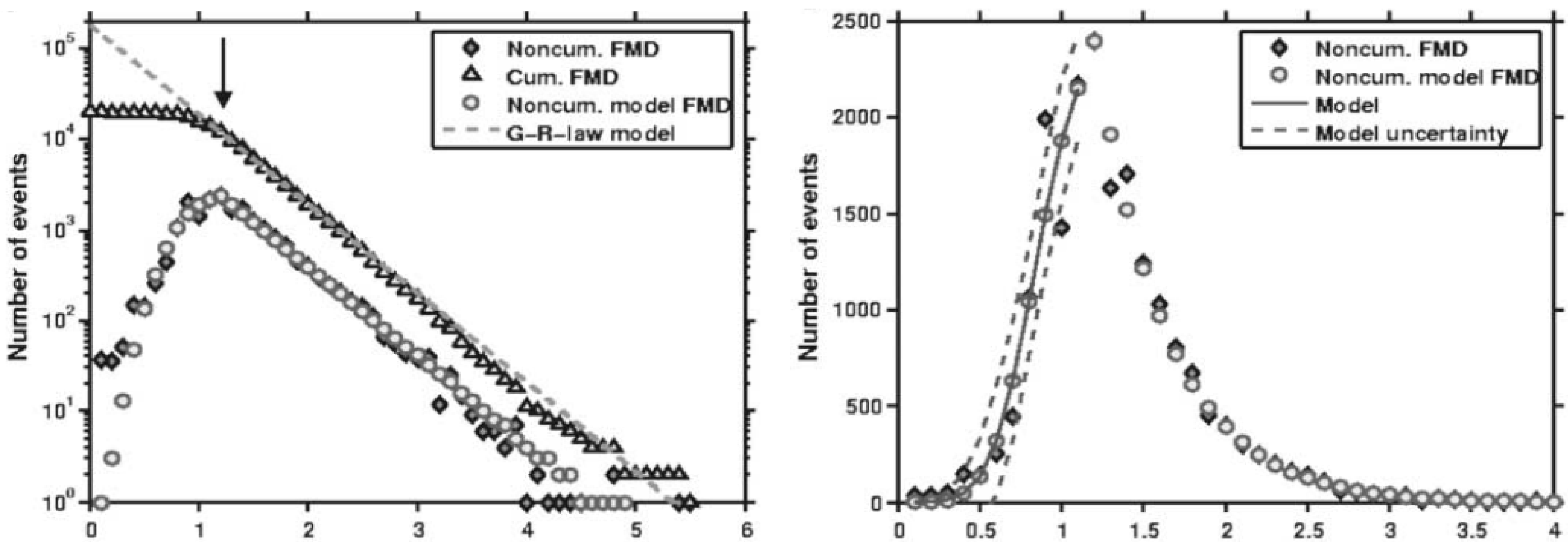
b-Value Stability Test

- Firstly proposed by Cao and Gao (2002)
- Later modified by Woessner and Wiemer (2005)
  - $\Delta b_i = \left|\bar{b}_i - b_i\right| \leq \delta b_i$
    - b_i: estimate of b-value for magnitude $m_{co} = m_i$
  - $\bar{b}_i = \frac{\sum_{k=i}^{i+K-1} b_i}{K} \leftarrow K=5$, $\bar{b}_i$ is quite sensitive to K
  - $\delta b_i = 2.3{b_i}^2\sqrt{\frac{\sum_{n=i}^{N}(m_n - \bar{m}_i)^2}{(N-i+1)(N-i)}}$ (Shi & Bolt, 1982)
    - $\bar{m}_i = \frac{\sum_{n=i}^{N} m_n}{(N-i+1)}$



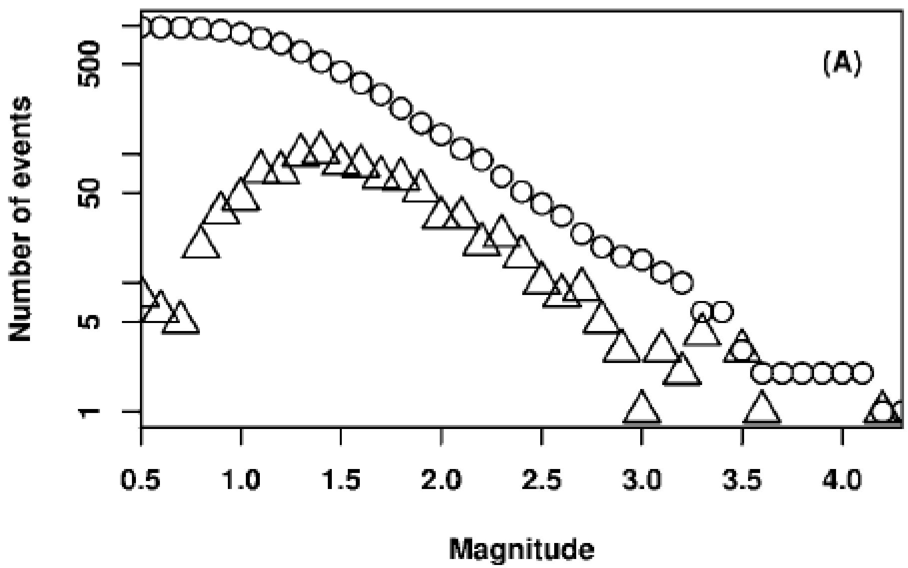
▣ Entire-Magnitude-Range Method

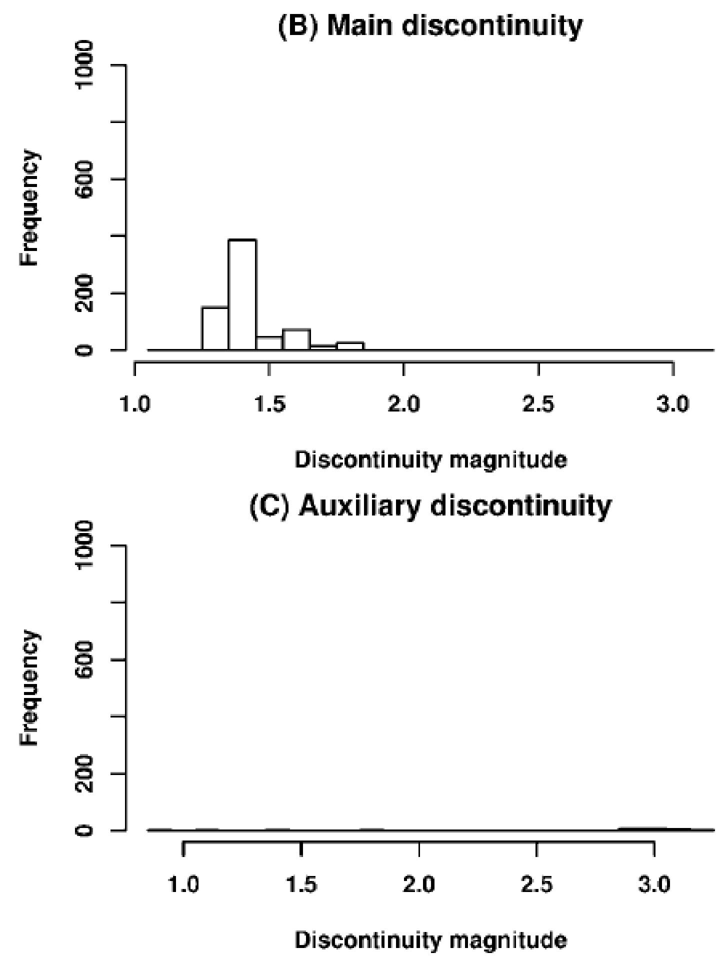
- ❖ Firstly proposed by Ogata and Katsura (1993)
- ❖ Later modified by Woessner and Wiemer (2005)
  - ➢ Maximum likelihood estimation of parameters
    - Modelling incomplete portion at smaller magnitudes by the error function
    - Modelling complete portion by exponential magnitude distribution
  - ➢ m_c to maximize sum of likelihoods for the two portions



▣ Change-Point Detection Method

- Amorese (2007)
- Detecting multiple change-points in b-estimates
- m_c to minimized the Type I error





▣ Chi-Square Test

❖ Noh (2019)

❖ Pearson's test statistic: $PTS(l) = \sum_{i=l}^{L} \frac{\left(n_i^{obs} - n_i^{pre}\right)^2}{n_i^{pre}}$

➢ n_i^{obs}: number of observed events with $m_i - \frac{\Delta m}{2} \le m < m_i + \frac{\Delta m}{2}$

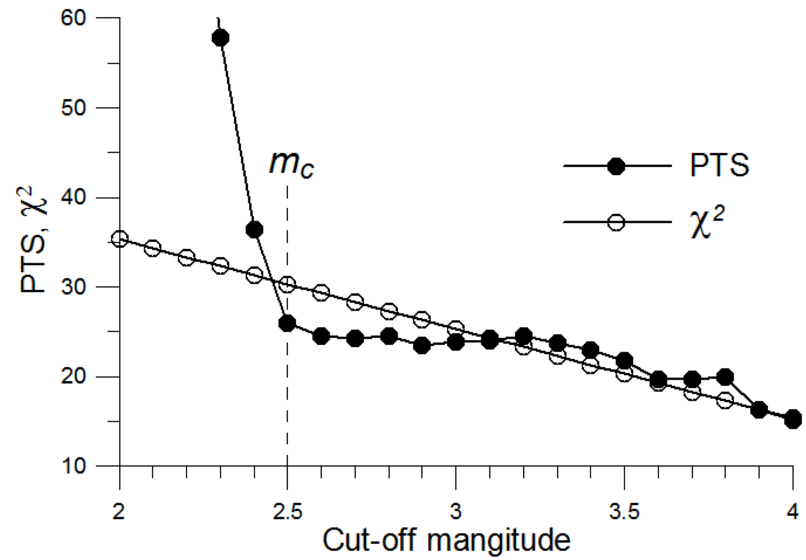
➢ n_i^{pre} : number of predicted events with $m_l - \frac{\Delta m}{2} \le m < m_i + \frac{\Delta m}{2}$

- $n_i^{pre} = p_{0i}\, n^{obs}$
- $p_{0i} = \frac{e^{-\beta m_i}}{\sum_{j=l}^{L} e^{-\beta m_j}}$ and $n^{obs} = \sum_{j=1}^{L} n_i^{obs}$

➢ $PTS(l) \sim \chi^2(L - l - 2)$

- Three constraints

❖ M_c : 1st cross-over magnitude



Chapter 6
Characterization of Seismic Sources
- Catalog Based -

Introduction

▣ Major Seismicity Parameters

- ❖ Maximum magnitude, Richter-b, annual occurrence rate

▣ Seismic vs. Geologic Approaches

- ❖ Seismicity-Based Approaches (Probabilistic)
  - ➢ Open the only option in regions with limited seismic record and limited geological investigations
  - ➢ Particularly useful for constraining rates of small to moderate events that do not provide surface evidence
- ❖ Geological Approaches (Deterministic)
  - ➢ Works well in active areas with a significant history of earthquake occurrence and geological investigations
  - ➢ Particularly useful for constraining rates of the largest events with surface evidences
- ❖ Cross Check
  - ➢ If two approaches are available, their estimates can be used for the cross check

▣ Inadequacy of LSM

❖ Common Assumptions

- Almost always
  - Independency of samples (i.e., observed data)
- In most cases
  - Independent, identically distributed (i.i.d. assumption)

❖ Least-Squares Method (MSM)

- Log-linear fitting of G-R relation
  - $\log N = a - bM$, where N is the number of events $\geq M$
- Violation of independency assumption
  - A change of the frequency at a magnitude affects all frequencies at magnitudes less than that magnitude
- Larger events are repeatedly counted in the smaller event counts
  - Lower b-values (Bender, 1983)

Magnitude Distribution

▣ Exponential Model

❖ Gutenberg-Richter Relation

- $\log N = a - bm \rightarrow N = 10^{a-bm}$
- For $m \geq m_0$, $N = N_0 e^{-\beta(m-m_0)}$
  - $N_0 = 10^{a-bm_0} = e^{\alpha-\beta m_0}$, $\alpha = a \ln 10$, and $\beta = b \ln 10$

❖ Derivation of PDF for $m_{max} \rightarrow \infty$

- $f_M(m)dm = \frac{k'[-dN(m)]}{N_0} = -\frac{k' \frac{dN(m)}{dm} dm}{N_0} = k'\beta e^{-\beta(m-m_0)} dm$
- Normalization:
  - $\int_{m_0}^{\infty} f_M(m)dm = k'\beta \int_{m_0}^{\infty} e^{-\beta(m-m_0)} dm = -k' e^{-\beta(m-m_0)} \Big|_{m_0}^{\infty} = k' = 1$
- PDF: $f_M(m) = \begin{cases} 0 & , m < m_0 \\ \beta e^{-\beta(m-m_0)} & , m_0 \leq m \end{cases}$
- CDF: $F_M(m) = \begin{cases} 0 & , m < m_0 \\ 1 - e^{-\beta(m-m_0)} & , m_0 \leq m \end{cases}$

▣ Exponential Model (continued)

❖ Introducing the magnitude upper bound m_{max}

- $1 = k[F_M(m_{max}) - F_M(m_0)] = k\left[1 - e^{-\beta(m_{max}-m_0)}\right]$ or
- $k = \left[1 - e^{-\beta(m_{max}-m_0)}\right]^{-1}$

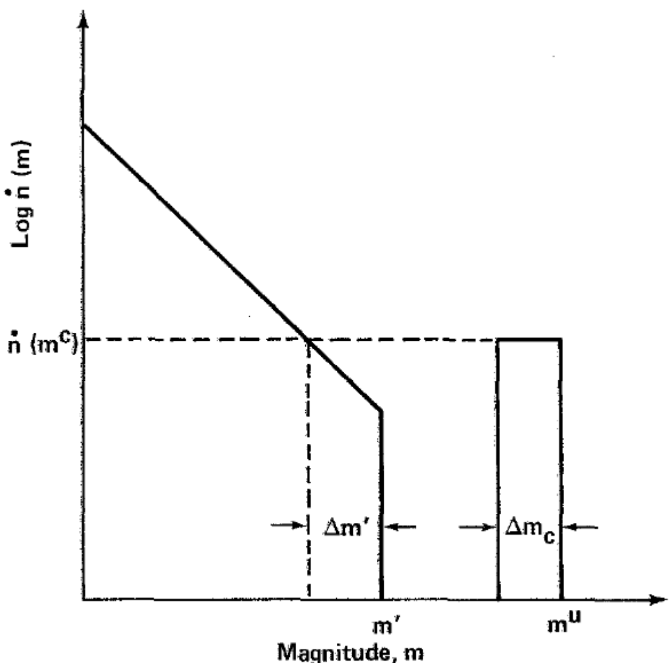
- PDF: $f_M(m) = \begin{cases} 0 & , m < m_0 \\ \dfrac{\beta e^{-\beta(m-m_0)}}{1-e^{-\beta(m_{max}-m_0)}} & , m_0 \le m \le m_{max} \\ 0 & m_{max} < m \end{cases}$

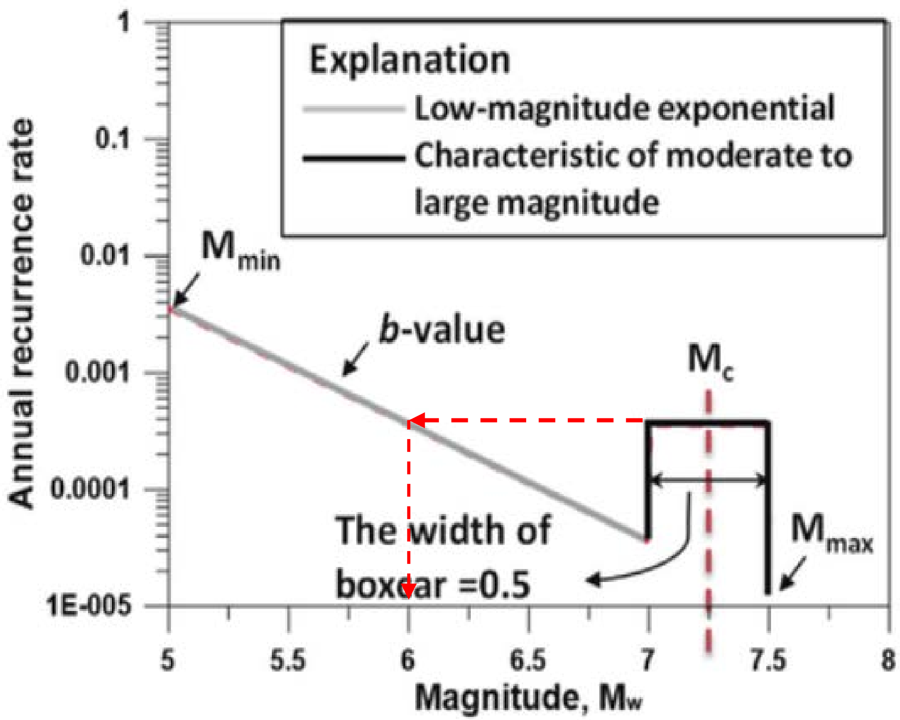
- CDF: $F_M(m) = \begin{cases} 0 & , m < m_0 \\ \dfrac{1-e^{-\beta(m-m_0)}}{1-e^{-\beta(m_{max}-m_0)}} & , m_0 \le m \le m_{max} \\ 1 & m_{max} < m \end{cases}$

▣ Characteristic Earthquake Model

❖ Schwartz and Coppersmith (1984)

- $\Delta m_c = 1/2, \quad m' = m^u - \Delta m_c$
- $\dot{n}^c = \dot{n}(m^c) = \dot{n}(m' - 1) \leftarrow \Delta m' = 1$





▣ Characteristic Earthquake Model (continued)

❖ PDF: $f_M(m) = \begin{cases} k'\beta e^{-\beta(m-m^0)}, & m^0 \le m \le m^u - 1/2 \\ k'\beta e^{-\beta(m^u-m^0-3/2)}, & m^u - 1/2 \le m \le m^u \\ 0, & \text{otherwise} \end{cases}$

where $q = \frac{1}{2}\frac{\beta e^{-\beta(m^u-m^0-3/2)}}{1-e^{-\beta(m^u-m^0-1/2)}}$ and $k' = \left[(1+q)\left(1-e^{-\beta(m^u-m^0-1/2)}\right)\right]^{-1}$

❖ Task: Derive the following formula

$$F_M(m) = \begin{cases} k'\left[1-e^{-\beta(m-m^0)}\right], & m^0 \le m \le m^u - 1/2 \\ k'\left[1-e^{-\beta\left(m^u-m^0-\frac{1}{2}\right)} + \beta e^{-\beta\left(m^u-m^0-\frac{3}{2}\right)}\left(m-m^u+\frac{1}{2}\right)\right], & m^u - 1/2 \le m \le m^u \\ 1, & m > m^u \end{cases}$$

Estimation of Richter-b

▣ Maximum likelihood method (MLM)

❖ Probability density function of magnitude

➢ $f_M(m) = k\beta \exp[-\beta(m - m_{min})]$

where $k^{-1} = 1 - \exp[-\beta(m_{max} - m_{min})]$, $\beta = b \ln 10$

➢ The parameter a has disappeared due to normalization of a PDF!

- Annual rate cannot be estimated from magnitude PDF only

❖ Likelihood function

➢ $L = \prod_{i=1}^{N} f_M(m_i) = (k\beta)^N \exp[-\beta \sum_{i=1}^{N}(m_i - m_{max})]$, or

➢ $\ln L = N \ln(k\beta) - \beta \sum_{i=1}^{N}(m_i - m_{min})$

$= N\,[\ln(k\beta) - \beta(\overline{m} - m_{min})]$, where $\overline{m} = \frac{\sum m_i}{N}$

❖ Maximum likelihood estimate: $\frac{\partial}{\partial \beta} \ln L = 0$ and $\frac{\partial^2}{\partial \beta^2} \ln L < 0$

➢ For $m_{max} \to \infty$; $k \to 1$

- $\frac{1}{\beta} = \overline{m} - m_{min}$ (Aki, 1965); (Utsu, 1965) by the method of moments
- $\frac{1}{\beta} \frac{\beta\delta}{\tanh(\beta\delta)} = \overline{m} - m_{min}$ for magnitude grouping with 2δ (Utsu, 1966)

▣ Maximum likelihood method (MLM)

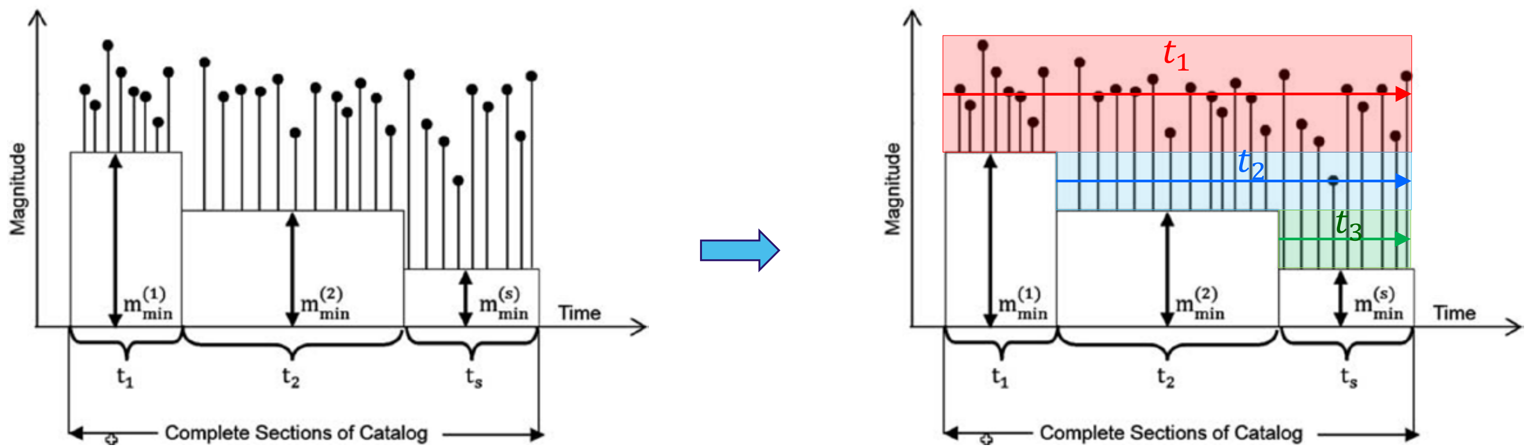
❖ Maximum likelihood estimate (continued)

➤ For finite m_{max}

- $\frac{1}{\widehat{\beta}} = \bar{m} - \frac{m_{min} - m_{max} \exp[-\beta(m_{max} - m_{min})]}{1 - \exp[-\beta(m_{max} - m_{min})]}$ (Page, 1968)
- $\bar{m} - \frac{m_{max} + m_{min}}{2} = \frac{1}{\widehat{\beta}} \left[\frac{\widehat{\beta}\delta}{\tanh(\widehat{\beta}\delta)} - \frac{\widehat{\beta}\frac{m_{max} - m_{min}}{2}}{\tanh\left(\widehat{\beta}\frac{m_{max} - m_{min}}{2}\right)} \right]$ for magnitude grouping with 2δ (Karnik, 1971)
- $\frac{\sum t_i\, m_i \exp(-\widehat{\beta} m_i)}{\sum t_i \exp(-\widehat{\beta} m_i)} = \frac{\sum n_i\, m_i}{N} = \bar{m}$ for magnitude grouping with 2δ & unequal observation period, t_i (Weichert, 1980), using $p(m_i) = P(m_i - \delta \le m < m_i + \delta) = \frac{t_i\, e^{-\beta m_i}}{\sum_{j=1}^{J} t_j\, e^{-\beta m_j}}$

- Extension to incomplete catalogs (Kijko & Smit, 2012)
  - $\hat{\beta} = \left(\frac{r_1}{\hat{\beta}_1} + \frac{r_2}{\hat{\beta}_2} + \cdots \frac{r_s}{\hat{\beta}_s}\right)^{-1}$

    where $r_i = n_i/n$ and $n = \sum_{i=1}^{s} n_i$



Estimation of Annual Rate

▣ Basic Approach

- ❖ For N events during T years
  - ➢ if n_k is the annual rate of events in k-th year
    - ▪ $\sum_{k=1}^{T} n_k = N$
- ❖ For the Poisson process with mean annual rate, ν
  - ➢ $P_N(n_k) = \frac{(\nu)^{n_k} e^{-\nu}}{n_k!}$
- ❖ Likelihood function
  - ➢ $L(\nu) = \prod_{k=1}^{T} P_N(n_k) = e^{-T\nu} \prod_{k=1}^{T} \frac{(\nu)^{n_k}}{n_k!}$, or
  - ➢ $\ln L(\nu) = -T\nu + \sum_{k=1}^{T}(n_k \ln \nu - \ln n_k!) = -T\nu + N \ln \nu - \sum_{k=1}^{T} \ln n_k!$
- ❖ ML Solution
  - ➢ $\frac{\partial \ln L}{\partial \nu} = 0$, or $\hat{\nu} = \frac{N}{T}$
  - ➢ $Var(\hat{\nu}) = -\left[\frac{\partial^2}{\partial \hat{\nu}^2} \ln L\right]^{-1}\Bigg|_{\nu=\hat{\nu}} = \frac{\hat{\nu}^2}{N}$

▣ Refined Formulation

- ❖ Exponential distribution with $m_{max} \to \infty$
- ❖ Mean frequency ρ_i of events of magnitude $(m_i, m_i + dm)$
  - ➢ $\rho_i = T\nu f_M(m_i)dm = T\nu\beta e^{-\beta(m_i - m_{min})}dm,\ m_i > m_{min}$
  - ➢ $P_N(n_i) = \dfrac{(\rho_i)^{n_i} e^{-\rho_i}}{n_i!}$
- ❖ For such a small dm that no more than one event in any magnitude interval
  - ➢ $P_N(n_i) = \begin{pmatrix} e^{-\rho_i}, & \text{if } no \text{ event},\ n_i = 0 \\ \rho_i\, e^{-\rho_i}, & \text{if } one \text{ event},\ n_i = 1 \end{pmatrix}$, $i = 1, 2, \cdots, I$, or
  - ➢ $P_N(n_i) = \begin{pmatrix} \exp\left[-\nu T\beta e^{-\beta(m_i - m_{min})}dm\right], & \text{if } no \text{ event} \\ \nu T\beta e^{-\beta(m_i - m_{min})}dm \times \exp\left[-\nu T\beta e^{-\beta(m_i - m_{min})}dm\right], & \text{if } one \text{ event} \end{pmatrix}$
- ❖ Likelihood function, as $dm \to 0$ $(I \to \infty)$
  - ➢ $L(\nu, \beta) = \lim_{I\to\infty}\left[\prod_{i=1}^{I} P_N(n_i)\right] = \lim_{I\to\infty}\left[\left(\prod_{i\neq j} e^{-\rho_i}\right) \times \left(\prod_{j=1}^{N} \rho_j e^{-\rho_j}\right)\right]$

$$= \left(\prod_{j=1}^{N} \rho_i\right) \times \left\{\lim_{I\to\infty}\left[\exp\left(-\sum_{i=1}^{I} \rho_i\right)\right]\right\}$$

$$\approx \prod_{i=1}^{N}\left[\nu T\beta e^{-\beta(m_i - m_{min})}dm\right] \times \exp\left[-\int_{m_{min}}^{\infty} \nu T\beta e^{-\beta(m - m_{min})}dm\right]$$

$= (\nu T dm)^N e^{-\nu T} \times \prod_{i=1}^{N}\left[\beta e^{-\beta(m_i - m_{min})}\right]$, or

Refined Formulation (continued)

- $\ln L = N \ln(\nu T dm) - \nu T + \sum_{i=1}^{N} [\ln \beta - \beta(m_i - m_{min})]$

ML Solution

- $\hat{\nu} = \frac{N}{T}$
- $\frac{1}{\hat{\beta}} = \frac{1}{N} \sum_{i=1}^{N} (m_i - m_{min}) = \bar{m} - m_{min}$
- Estimation of $\hat{\nu}$ and $\hat{\beta}$ is completely separated!

Tasks

1. Calculate variance of the above estimate of $\hat{\nu}$.

[Hint] Use $Var(\hat{\nu}) = -\left[\frac{\partial^2}{\partial \hat{\nu}^2} \ln L\right]^{-1}\Bigg|_{\nu=\hat{\nu}}$.

2. Extend estimate of $\hat{\nu}$ for a finite m_{max}.

[Hint] Replace $f_M(m_i) = \beta e^{-\beta(m_i - m_{min})}$ by $f_M(m_i) = k\beta e^{-\beta(m_i - m_{min})}$, where $k = \left(1 - e^{-\beta(m_{max} - m_{min})}\right)^{-1}$ and the upper integration limit ∞ by m_{max}.

▣ Magnitude-Grouped, Unequal Observation Time

❖ Noh (unpublished)

❖ Probability over i-th magnitude interval $(m_i, m_i + dm)$

➢ $p_i = p(m_i) = P(m_i - \delta \le m < m_i + \delta) = \frac{e^{-\beta m_i}}{\sum_{k=1}^{I} e^{-\beta m_k}}$

$\leftrightarrow p'_i = \frac{t_i e^{-\beta m_i}}{\sum_{k=1}^{I} t_k e^{-\beta m_k}}$ (Weichert, 1980)

❖ Mean frequency ρ_i of events of i-th magnitude interval

➢ $\lambda_i = \nu t_i p_i = \frac{\nu t_i e^{-\beta m_i}}{\sum_{k=1}^{I} e^{-\beta m_k}}$

❖ Poisson probability for frequency n_i

➢ $P_N(n_i) = \frac{(\lambda_i)^{n_i} e^{-\lambda_i}}{n_i!} = \frac{(\nu t_i p_i)^{n_i} e^{-\nu t_i p_i}}{n_i!}$

❖ Log-likelihood function

➢ $\ln L(\nu, \beta) = \sum_{i=1}^{I} \ln[P_N(n_i)] = \sum_{i=1}^{I} [n_i \ln(\lambda_i) - \lambda_i - \ln(n_i!)]$

❖ Log-likelihood function (continued)

- $\ln L(\nu,\beta) = \sum_{i=1}^{I} \ln[P_N(n_i)] = \sum_{i=1}^{I}[n_i \ln(\lambda_i) - \lambda_i - \ln(n_i!)]$

  $= \sum_{i=1}^{I}\left[n_i \ln\nu + n_i \ln t_i - n_i\beta m_i - n_i \ln S - \nu t_i e^{-\beta m_i}/S - \ln(n_i!)\right]$

  $= N\ln\nu + \sum_{i=1}^{I}(n_i \ln t_i) - \beta N\bar{m} - N\ln S - \nu S_t/S - \ln(n_i!)$

  where $N = \sum_{i=1}^{I} n_i$, $\bar{m} = \frac{1}{N}\sum_{i=1}^{I} n_i m_i$, $S = \sum_{k=1}^{I} e^{-\beta m_k}$, and $S_t = \sum_{k=1}^{I} t_k\, e^{-\beta m_k}$

❖ Estimation of ν

- $\frac{\partial}{\partial\nu}\ln L = \frac{N}{\nu} - \frac{S_t}{S}$

- $\hat{\nu} = \frac{NS}{S_t} = \frac{\sum_{i=1}^{I} e^{-\hat{\beta}m_i}}{\sum_{k=1}^{I} t_k e^{-\hat{\beta}m_k}} N$ (1)

- $Var(\hat{\nu}) = -\left[\frac{\partial^2}{\partial\hat{\nu}^2}\ln L\right]^{-1}\Bigg|_{\nu=\hat{\nu}} = \frac{\hat{\nu}^2}{N}$ (2)

  - Estimation of $\hat{\nu}$ and $\hat{\beta}$ is not separated!

- $\hat{\nu}_{m\geq m_l} = \hat{\nu}\frac{\sum_{k=l}^{I} e^{-\hat{\beta}m_k}}{\sum_{k=1}^{I} e^{-\hat{\beta}m_k}} = \frac{\hat{\nu}}{S}\sum_{k=l}^{I} e^{-\hat{\beta}m_k}$

❖ Estimation of β

➢ $$\frac{\partial}{\partial\beta}\ln L = -N\bar{m} + \frac{NS_m}{S} + \nu\frac{S_{tm}S - S_t S_m}{S^2}, \qquad \because S_m = \sum_{k=1}^{I} m_k\, e^{-\beta m_k}$$

$$= -N\bar{m} + \frac{NS_m}{S} + \nu\frac{S_{tm}S - S_t S_m}{S^2}, \qquad \because S_{tm} = \sum_{k=1}^{I} t_k m_k e^{-\beta m_k}$$

$$= -N\bar{m} + \frac{NS_m}{S} + \frac{NS}{S_t}\,\frac{S_{tm}S - S_t S_m}{S^2}, \qquad \because \nu = \frac{NS}{S_t}$$

$$= -N\bar{m} + \frac{NS_m}{S} + \frac{NS_{tm}}{S_t} - \frac{NS_m}{S}$$

$$= -N\left(\bar{m} - \frac{S_{tm}}{S_t}\right)$$

➢ $$\bar{m} = \frac{S_{tm}}{S_t} = \frac{\sum_{i=1}^{I} t_i\, m_i\, e^{-\hat{\beta} m_i}}{\sum_{k=1}^{I} t_k\, e^{-\hat{\beta} m_k}} \tag{3}$$

➢ $$Var(\hat{\beta}) = -\left[\frac{\partial^2}{\partial\beta^2}\ln L\right]^{-1}\Bigg|_{\beta=\hat{\beta}} = \frac{1}{N}\frac{S_t^2}{S_{tm}^2 - S_{tmm}S_t}\Bigg|_{\beta=\hat{\beta}} \tag{4}$$

where $S_{tmm} = \sum_{k=1}^{I} t_k m_k m_k e^{-\beta m_k}$

Estimation of m_{max}

▣ Introduction

❖ Why no maximum likelihood estimates using $f_M(m)$?

➢ $\ln L = n\ln(k\beta) - \beta\sum_{i=1}^{n}(m_i - m_{min})$

$= n\,[\ln(k\beta) - \beta(\bar{m} - m_{min})],$

where $k^{-1} = 1 - \exp[-\beta(m_{max} - m_{min})] > 0$

➢ $\frac{\partial \ln L}{\partial m_{max}} = -\frac{n\beta e^{-\beta(m_{max}-m_{min})}}{1-e^{-\beta(m_{max}-m_{min})}} < 0$

❖ General form of m_{max} estimator

➢ $m_{max} = m_{max}^{obs} + \Delta_n$

➢ Usually, Δ_n includes m_{max}

- $\Delta_n = \int_{m_{min}}^{m_{max}}\left[\frac{1-\exp[-\beta(m-m_{min})]}{1-\exp[-\beta(m_{max}-m_{min})]}\right]^n$ (Kijko, 2004)

➢ (inner) Iteration scheme is required

➢ $Var(\hat{m}_{max}) = \sigma_{M_x^o}^2 + \sigma_M^2$

- $\sigma_{M_x^o}^2$: uncertainty related to the determination of m_{max}^{obs}
- σ_M^2 : uncertainty related to the magnitude determination ($\cong {\Delta_n}^2$)

❖ List of Methods

Class	Name	Remark
Parametric	T-P	Procedure by Pisarenko et al. (1996)
	K-S	Procedure by Kijko & Sellevoll (1989)
	T-P-B	Tate-Pisarenko-Bayes procedure
	K-S-B	Kijko-Sellevoll-Bayes procedure
Non-parametric	N-P-G	Non-parametric procedure with Gaussian kernel
	N-P-OS	Non-parametric procedure based on order statistics
	R-W	Robson-Whitlock procedure
	R-W-C	Robson-Whitlock-Cooke procedure
	F-L-E	Procedure based on a few large earthquakes
Fit of CDF	L1-Fit	Procedure based on fit of L1 norm CDF
	L2-Fit	Procedure based on fit of L2 norm CDF

▣ Parametric Approaches

❖ Tate-Pisarenko Procedure

- Order statistics of earthquake magnitude: $M_1 \le M_2 \le \cdots \le M_n$
  - M_i is independent, identically distributed by $F_M(m|m_{max})$
- For transformation $Y_i = F_M(M_i|m_{max})$
  - Y_i is a uniform deviate such that

    $Y_1 \le Y_2 \le \cdots \le Y_n$ and

$$F_Y(y) = \begin{cases} 0, & y < 0 \\ y, & 0 \le y \le 1 \\ 1, & y > 1 \end{cases}$$

- CDF of the largest among Y_i, that is Y_n is
  - $F_{Y_n}(y) = P[Y_n \le y] = P[Y_1 \le y, Y_2 \le y, \cdots, Y_n \le y]$

    $= [F_Y(y)]^n = y^n$
- PDF of Y_n is
  - $f_{Y_n}(y) = \begin{cases} 0, & y < 0 \\ ny^{n-1}, & 0 \le y \le 1 \\ 0, & y > 1 \end{cases}$

❖ Tate-Pisarenko Procedure (continued)

- Expectation
  - $E(Y_n) = \int_0^1 \xi f_{Y_n}(\xi) d\xi = n \int_0^1 \xi^n d\xi = \frac{n}{n+1}$ (1)
- Best unbiased estimation of $E(Y_n)$ is y_n
  - $E(Y_n) = y_n$ (2)
- Using the Taylor expansion of $M_n = F_M^{-1}(Y_n|m_{max})$ at $Y_n = 1$
  - $M_n = F_M^{-1}(1|m_{max}) - \left.\frac{dF_M^{-1}(Y_n|m_{max})}{dY_n}\right|_{Y_n=1} (1 - Y_n) + \cdots$ (3)
- Taking average of both sides of (3) and using
  - $E(M_n) = m_{max}^{obs}$
  - $F_M^{-1}(1|m_{max}) = m_{max}$
  - $E(1 - Y_n) = 1 - \frac{n}{n+1} = \frac{1}{n+1}$
  - $\left.\frac{dF_M^{-1}(Y_n|m_{max})}{dY_n}\right|_{Y_n=1} = \frac{1}{\left.\frac{dF_M(M_n|m_{max})}{dM_n}\right|_{M_n=m_{max}}} = \frac{1}{f_M(m_{max}|m_{max})}$
- We arrive at
  - $m_{max}^{obs} = m_{max} - \frac{1}{(n+1)f_M(m_{max}|m_{max})}$

❖ Tate-Pisarenko Procedure (continued)

- For a large n
  - $E(1 - Y_n) = \frac{1}{n+1} \cong \frac{1}{n}$
  - $f_M(m_{max}|m_{max}) \cong f_M(m_{max}^{obs}|m_{max}^{obs})$
- Finally
  - $m_{max} = m_{max}^{obs} + \frac{1}{n f_M(m_{max}^{obs}|m_{max}^{obs})}$
- $\Delta_n = \frac{1}{n f_M(m_{max}^{obs}|m_{max}^{obs})}$
- For doubly truncated PDF,
  - $\Delta_n = \frac{1-\exp[-\beta(m_{max}^{obs}-m_{min})]}{n\beta \exp[-\beta(m_{max}^{obs}-m_{min})]}$
- The estimator is,
  - $m_{max} = m_{max}^{obs} + \frac{1-\exp[-\beta(m_{max}^{obs}-m_{min})]}{n\beta \exp[-\beta(m_{max}^{obs}-m_{min})]}$
  - $Var(\widehat{m}_{max}) = \sigma_{M_x^o}^2 + \Delta_n^2$

❖ Notes

- (5) was probably first derived by Tate (1959)
- It was used by Pisarenko *et al.* (1996)

❖ Kijko-Sellevoll Procedure

- Kijko & Sellevoll (1989)
- From order statistics, CDF of the largest observed magnitude among n events, $m_n \equiv m_{max}^{obs}$ is $F_{M_n}(m) = [F_M(m)]^n$
  - $E(M_n) = \int_{m_{min}}^{m_{max}} m dF_{M_n}(m) = m_{max} - \int_{m_{min}}^{m_{max}} F_{M_n}(m)dm$ or
  - $m_{max} = E(M_n) + \int_{m_{min}}^{m_{max}} F_{M_n}(m)dm$ or
  - $m_{max} = m_{max}^{obs} + \int_{m_{min}}^{m_{max}} [F_M(m)]^n dm$
- For large n, $[F_M(m)]^n \approx \exp\{-n[1 - F_M(m)]\}$ (Cramér, 1961)
- For doubly truncated PDF,
  - $\Delta_n \approx \int_{m_{min}}^{m_{max}} \exp\{-n[1 - F_M(m)]\} dm = \frac{E_1(n_2) - E_1(n_1)}{\beta \exp(-n_2)} + m_{min}\exp(-n)$
    - $n_1 = \frac{n}{\{1-\exp[-\beta(m_{max}-m_{min})]\}}$, $n_2 = n_1 \exp[-\beta(m_{max} - m_{min})]$, and
    - $E_1(z) = \int_z^\infty \frac{\exp(-\omega)}{\omega} d\omega$; exponential integration function

❖ Kijko-Sellevoll Procedure (continued)

- The estimator is,
  - $m_{max} = m_{max}^{obs} + \frac{E_1(n_2) - E_1(n_1)}{\beta \exp(-n_2)} + m_{min} \exp(-n)$
  - $Var(\widehat{m}_{max}) = \sigma_{M_x^o}^2 + {\Delta_n}^2$
- While the exact formula of Δ_n is reported, it is not discussed here because it does not gives an improved accuracy but is just complicated.

※ A direct numerical integration, such as the Romberg integration, of $\Delta_n = \int_{m_{min}}^{m_{max}} [F_M(m)]^n dm$ yields an accurate enough result.

❖ Tate-Pisarenko-Bayes Procedure

- Assuming a gamma distribution for $f_{\mathrm{B}}(\beta)$, Campbell (1982) showed
  - $$f_M(m) = \begin{cases} 0 & , m < m_{min} \\ \bar{\beta} C_\beta \left(\frac{p}{p+m-m_{min}}\right)^{q+1} & , m_{min} \le m \le m_{max} \\ 0 & , m > m_{max} \end{cases}$$
  - $$F_M(m) = \begin{cases} 0 & , m < m_{min} \\ C_\beta \left[1 - \left(\frac{p}{p+m-m_{min}}\right)^{q}\right] & , m_{min} \le m \le m_{max} \\ 0 & , m > m_{max} \end{cases}$$
    - $C_\beta = \left\{1 - \left(\frac{p}{p+m_{max}-m_{min}}\right)^{q}\right\}^{-1}$, $p = \frac{\bar{\beta}}{\sigma_\beta^2}$, $q = \left(\frac{\bar{\beta}}{\sigma_\beta}\right)^2$
    - $\bar{\beta}$ is a known value of β and σ_β a known standard deviation of β, of which values are taken from the their estimates to be discussed in the subsequent section
- For doubly truncated PDF,
  - $\Delta_n = \frac{1}{n\bar{\beta}C_\beta}\left(\frac{p}{p+m_{obs}-m_{min}}\right)^{-(q+1)}$
  - $m_{max} = m_{max}^{obs} + \Delta_n$
  - $Var(\hat{m}_{max}) = \sigma_{M_x^o}^2 + {\Delta_n}^2$
- T-P-B yields estimate of m_{max} very close to that of T-P

❖ Kijko-Sellevoll-Bayes Procedure

- Assuming a gamma distribution for $f_{\mathrm{B}}(\beta)$, Campbell (1982)
  - $\Delta_n = \left(C_\beta\right)^n \int_{m_{min}}^{m_{max}} \left[1 - \left(\frac{p}{p+m-m_{min}}\right)^q\right]^n dm$
- Using Cramér's approximation
  - $\Delta_n = \frac{\delta^{1/q} \exp[nr^q/(1-r^q)]}{\bar{\beta}} \left[\Gamma\left(-\frac{1}{q}, \delta r^q\right) - \Gamma\left(-\frac{1}{q}, \delta\right)\right]$,

    where $r = p/(p + m_{max} - m_{min})$, $\delta = nC_\beta$
  - $m_{max} = m_{max}^{obs} + \Delta_n$
  - $Var(\hat{m}_{max}) = \sigma_{M_x^o}^2 + {\Delta_n}^2$
- K-S-B yields estimate of m_{max} very close to that of K-S

▣ Non-Parametric Approaches

❖ Non-Parametric with Gaussian Kernel Procedure

➢ Kernel estimator $\hat{f}_M(m)$ of actual, unknown PDF $f_M(m)$

- $\hat{f}_M(m) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{m-m_i}{h}\right)$
  - h : positive smoothing factor
  - $K(\cdot)$: kernel function, a PDF, symmetric about zero
- Estimation is not sensitive to the kernel function
  - Choice is the standard normal PDF, $K(z) = (2\pi)^{-1/2}\exp(-z^2/2)$ normalized in the range $\left[\frac{m_{min}-m_i}{h}, \frac{m_{max}-m_i}{h}\right]$
  - But the choice of a smoothing factor is crucial
- $$\hat{f}_M(m) = \begin{cases} 0 & , m < m_{min} \\ \frac{1}{\sqrt{2\pi}\, nh}\sum_{i=1}^{n} \frac{\exp\left[-\left(\frac{m-m_i}{\sqrt{2}h}\right)^2\right]}{\Phi\left(\frac{m_{max}-m_i}{h}\right)-\Phi\left(\frac{m_{min}-m_i}{h}\right)} & , m_{min} \le m \le m_{max} \\ 0 & , m > m_{max} \end{cases}$$
  - $\Phi(z)$: standard normal CDF

❖ Non-Parametric with Gaussian Kernel Procedure (continued)

- $$\hat{F}_M(m) = \begin{cases} 0 & , m < m_{min} \\ \frac{1}{n}\sum_{i=1}^{n} \frac{\Phi\left(\frac{m-m_i}{h}\right)-\Phi\left(\frac{m_{min}-m_i}{h}\right)}{\Phi\left(\frac{m_{max}-m_i}{h}\right)-\Phi\left(\frac{m_{min}-m_i}{h}\right)} & , m_{min} \le m \le m_{max} \\ 1 & , m > m_{max} \end{cases}$$

➢ Estimators

- $m_{max} = m_{max}^{obs} + \Delta_n$
- $Var(\widehat{m}_{max}) = \sigma_{M_x^o}^2 + \Delta_n^2$
- T-P procedure: $\Delta_n = \frac{1}{n\hat{f}_M(m_{max}^{obs})}$
- K-S procedure: $\Delta_n = \int_{m_{min}}^{m_{max}} \left[\hat{F}_M(m)\right]^n dm$

❖ Non-Parametric Procedure Based on Order Statistics

- For ordered n observations, $m_1 \le m_2 \le \cdots \le m_{n-1} \le m_n$
  - $\hat{F}_M(m) = \begin{cases} 0 & , m < m_1 \\ \frac{i}{n} & , m_i \le m \le m_{i+1} \\ 1 & , m > m_n \end{cases}$ (Cooke, 1979)
- Approximate of integral Δ_n
  - $\Delta_n \equiv \int_{m_{min}}^{m_{max}^{obs}} \left[\hat{F}_M(m)\right]^n = \sum_{i=1}^{n-1} \left(\frac{1}{n}\right)^n (m_{i+1} - m_i)$

    $= m_{max}^{obs} - \sum_{i=0}^{n-1} \left[\left(1 - \frac{i}{n}\right)^n - \left(1 - \frac{i+1}{n}\right)^n\right] m_{n-i}$
  - For large n, $(1 + 1/n)^n \cong e$

    $\Delta_n \cong m_{max}^{obs} - (1 - e^{-1}) \sum_{i=1}^{n-1} e^{-i} m_{n-i}$
- Estimator of m_{max}
  - $m_{max} = m_{max}^{obs} + \Delta_n$
  - $Var(\hat{m}_{max}) = c_0 \sigma_{M_x^o}^2 + \Delta_n^2$
    - $c_0 = (1 + e^{-1})^2 + e^{-2}(1 - e^{-1})/(1 + e^{-1}) \cong 1.93$

❖ Robson-Whitlock Procedure

- For ordered n observations, $m_1 \le m_2 \le \cdots \le m_{n-1} \le m_n$, Robson and Whitlock (1964) proposed
  - $\hat{m}_{max} = m_{max}^{obs} + \left(m_{max}^{obs} - m_{n-1}\right)$
- For a doubly-truncated exponential distribution
  - $Var(\hat{m}_{max}) = 5\sigma_{M_x^o}^2 + \Delta_n^2,\quad \Delta_n = m_{max}^{obs} - m_{n-1}$
- While its simplicity makes it very attractive, it is known that reduction of bias is achieved at the expense of mean squared error.

❖ Robson-Whitlock-Cooke Procedure

- Cooke (1979) showed that reduction of the mean squared error of the R-W estimator is possible when some information, ν about the shape of the upper tail of PDF, $f_M(m)$
  - $\hat{m}_{max} = m_{max}^{obs} + (2\nu)^{-1}\left(m_{max}^{obs} - m_{n-1}\right)$
- For a doubly-truncated exponential distribution, $\nu = 1$
  - $\hat{m}_{max} = m_{max}^{obs} + \frac{1}{2}\left(m_{max}^{obs} - m_{n-1}\right)$
  - $Var(\hat{m}_{max}) = \frac{3}{2}\sigma_{M_x^o}^2 + \Delta_n^2, \quad \Delta_n = \frac{1}{2}\left(m_{max}^{obs} - m_{n-1}\right)$

❖ Procedure Based on a Few Largest Earthquakes

- Gnedenko (1943) suggested for a very broad class of $F_M(m)$

  1) When m is near to the upper end point
  2) $F_M(m)$ is linear in m

  - $\hat{m}_{max} = \sum_{i=1}^{n_0} a_i \, m_{n-i+1}$
    - a_i : coefficients to be determined, $i = 1, \cdots, n_0$
    - m_k : order statistics
    - n_0 : the number of largest earthquakes

- For truncated distributions, the mean squared error of $\hat{m}_{max}$ is minimized when
  - $a_2 = \cdots = a_{n_o - 1} = 0$, and $a_{n_o} = -1/n_0$
  - That is, $\Delta_n = \frac{1}{n_0}\left(m_{max}^{obs} - m_{n-n_0+1}\right)$
  - Due to Quenouille (1965), an improved Δ_n is
    - $\Delta_n = \frac{1}{n_0}\left(m_{max}^{obs} - \frac{1}{n_0 - 1}\sum_{i=2}^{n_0} m_{n-i+1}\right)$

❖ Procedure Based on a Few Largest Earthquakes (continued)

- The estimators for m_{max}
  - $\hat{m}_{max} = m_{max}^{obs} + \Delta_n, \quad \Delta_n = \frac{1}{n_0}\left(m_{max}^{obs} - \frac{1}{n_0-1}\sum_{i=2}^{n_0} m_{n-i+1}\right)$
  - $Var(\hat{m}_{max}) = c_0\sigma_{M_x^o}^2 + \Delta_n^2, \quad c_0 = (n_0^2 + n_0 - 1)/[n_0(n_0 - 1)]$
- Note that
  1) When we have sufficient sample, $n_0 \gg 1$, $\Delta_n \approx 0$
  2) Therefore, this estimator is useful only when we have limited information, a few large earthquakes

Fit of CDF Approach

▣ Fit of CDF Approaches

❖ Procedure Based on L1-Norm of CDF

- For ordered n observations, $m_1 \leq m_2 \leq \cdots \leq m_{n-1} \leq m_n$, the set of model parameters $\boldsymbol{\theta}$ can be found by minimizing the misfit function
  - $J(\boldsymbol{\theta}) = \sum_{i=1}^{n} \left| F_M(m_i) - \hat{F}_M(m_i) \right|$, $\hat{F}_M(m_i) = i/(n+1)$
- In case of the doubly-truncated exponential PDF, $\boldsymbol{\theta} = (\beta, m_{max})$
- $\boldsymbol{\theta}$ can be calculated by numerical methods, such as simplex method (Press et al, 1994)
- Note that, the misfit function of L_1 norm could have multiple extrema for more than one parameter

❖ Procedure Based on L2-Norm of CDF

- For ordered n observations, $m_1 \leq m_2 \leq \cdots \leq m_{n-1} \leq m_n$, the set of model parameters $\boldsymbol{\theta}$ can be found by minimizing the misfit function
  - $J(\boldsymbol{\theta}) = \sum_{i=1}^{n} \left[F_M(m_i) - \hat{F}_M(m_i)\right]^2$, $\hat{F}_M(m_i) = 1/(n+1)$
- Solving this the least-squares method is equivalent to the maximum likelihood method with the assumption that the distribution of the CDF residuals is of Gaussian

▣ Variance of θ for the Gaussian Procedure

❖ Generalized misfit function to be minimized

➢ $J(\boldsymbol{\theta}) = \sum_{i=1}^{n} |q_i|^p = \sum_{i=1}^{n} |y_i - g_i(\boldsymbol{\theta})|^p, \quad p \in [1, 2)$

- y_i : i-th observation
- g_i : model prediction for i-th observation
- q_i : prediction error or noise at i-th observation

❖ For the generalized Gaussian process

➢ $f(q|\mu, \kappa, \beta) = \frac{\beta}{2\kappa\Gamma(\frac{1}{\beta})} \exp\left[-\left(\frac{|q-\mu|}{\kappa}\right)^{\beta}\right]$

- μ : location parameter (= 0, assuming q_i has a zero mean)
- κ : scale parameter

❖ The covariance matrix is

➢ $\boldsymbol{C} = \begin{cases} \frac{\Gamma(\frac{2p-1}{\beta})\Gamma(\frac{1}{\beta})}{(p-1)^2\Gamma^2(\frac{p-1}{\beta})} \kappa^2 \boldsymbol{U}^{-1} & , p > 1 \\ \Gamma^2(1 + \frac{1}{\beta})\kappa^2 \boldsymbol{U}^{-1} & , p = 1 \end{cases}$, where $u_{ij} = \sum_{k=1}^{n} g_{k,i} g_{k,j}$; $g_{k,i} = \frac{\partial g_k}{\partial x_i}$

▣ Variance of θ for the Gaussian Procedure (continued)

❖ Ordinary Gaussian process; $\beta = 2$

➢ $$\boldsymbol{C}_G = \begin{cases} \dfrac{\Gamma(\frac{2p-1}{2})\Gamma(\frac{1}{2})}{(p-1)^2\Gamma^2(\frac{p-1}{2})}\kappa^2\boldsymbol{U}^{-1} & ,p > 1 \\ \Gamma^2(\frac{3}{2})\kappa^2\boldsymbol{U}^{-1} & ,p = 1 \end{cases}$$

❖ L_1 Norm: $p = 1$

➢ $\boldsymbol{C}_{G|p=1} = \Gamma^2\left(\frac{3}{2}\right)\kappa^2\boldsymbol{U}^{-1} = \frac{\pi}{4}\kappa^2\boldsymbol{U}^{-1};\ \kappa = \frac{1}{n}\sum_{i=1}^{n}|q_i| \quad \because\ \mu = 0$

❖ L_2 Norm: $p = 2$

➢ $\boldsymbol{C}_{G|p=2} = \frac{1}{2}\kappa^2\boldsymbol{U}^{-1}\ ;\ \kappa = \sqrt{2}\sqrt{\frac{1}{n}\sum_{i=1}^{n}(q_i)^2} \quad \because\ \mu = 0$

▣ Variance of θ for the Gaussian Procedure (continued)

❖ Finally, the matrix $\boldsymbol{U}$ is calculated as follows

➢ $q_i = \frac{i}{n+1} - F_M(m_i|\beta, m_{max}) = \frac{i}{n+1} - \frac{1-e^{-\beta(m_i-m_{min})}}{1-e^{-\beta(m_{max}-m_{min})}}$

➢ $g_{i,1} = \frac{\partial g_i}{\partial \beta} = \frac{(m_i-m_{min})(1-e_x)e_i-(m_i-m_{min})(1-e_i)e_x}{(1-e_x)^2}$

➢ $g_{i,2} = \frac{\partial g_i}{\partial m_{max}} = -\frac{\beta(1-e_i)e_x}{(1-e_x)^2}$

➢ $e_i = e^{-\beta(m_i-m_{min})}$; $e_x = e^{-\beta(m_{max}-m_{min})}$

❖ Therefore,

➢ $Var(\hat{\beta}) = \left(\boldsymbol{C}_{G|p}\right)_{11}$; $Var(\widehat{m}_{max}) = \left(\boldsymbol{C}_{G|p}\right)_{22}$ where $p = 1$ or $p = 2$

▣ Alternative Approach to Estimate Variances

- ❖ Method in the previous slides is quite general, but somewhat complicated
- ❖ Considering the sensitivity of $\widehat{m}_{max}$ to $\hat{\beta}$, it would be better to separately estimate β by a proper method, if exists.
- ❖ We do have such a method, Weichert (1980) discussed in the section 'Estimation of Richter-b'
- ❖ Moreover, use of Weichert (1980) is consistent with the other $\widehat{m}_{max}$ estimators introduced in this course
- ❖ In the following, we use $\hat{\beta}$ by Weichert so that there is only one parameter to be estimated, $\widehat{m}_{max}$
- ❖ As before, the cost or misfit function is
  - ➢ $J(\boldsymbol{\theta}) = J(m_{max}) = \sum_{i=1}^{n} |q_i|^p = \sum_{i=1}^{n} |y_i - g_i|^p$
  - ➢ $y_i = \frac{i}{n+1}$ and $g_i = F_M(m_i | m_{max}) = \frac{1-e_i}{1-e_x}$
    - ▪ $e_i = \exp[-\beta(m_i - m_{min})]$ and $e_x = \exp[-\beta(m_{max} - m_{min})]$

Alternative Approach to Estimate Variances (continued)

- L_1 Norm: $p = 1$
  - $J(m_{max}) = \sum_{i=1}^{n} |q_i| = \sum_{i=1}^{n} \operatorname{sgn}(q_i)(y_i - g_i)$
  - Minimization of cost (misfit) function
    - $0 = \frac{\partial J}{\partial m_{max}} = -\sum_{i=1}^{n} \operatorname{sgn}(q_i) \frac{\partial g_i}{\partial m_{max}} = \frac{\beta e_x}{(1-e_x)^2} \sum_{i=1}^{n} \operatorname{sgn}(q_i)\,(1 - e_i)$ or
    - $\sum_{i=1}^{n} \operatorname{sgn}(q_i)\,(1 - e_i) = 0$
    - Can be solved by a root-finding algorithm
  - $u_{ij} = u_{22} = \sum_{i=1}^{n} \left(g_{i,2}\right)^2 = \frac{(\beta e_x)^2}{(1-e_x)^4} \sum_{i=1}^{n} (1 - e_i)^2 = s^2$
  - $Var(\widehat{m}_{max}) = \frac{\pi}{4}\left(\frac{\kappa}{s}\right)^2$, where $\kappa = \frac{1}{n}\sum_{i=1}^{n} |q_i|$

▣ Alternative Approach to Estimate Variances (continued)

❖ L_2 Norm: $p = 2$

- $J(m_{max}) = \sum_{i=1}^{n}(y_i - g_i)^2$
- Minimization of cost (misfit) function
  - $0 = \frac{\partial J}{\partial m_{max}} = -2\sum_{i=1}^{n} q_i \frac{\partial g_i}{\partial m_{max}} = \frac{\beta e_x}{(1-e_x)^2}\sum_{i=1}^{n} q_i(1-e_i)$ or
  - $\sum_{i=1}^{n} q_i(1-e_i) = 0$
  - Can be solved by a root-finding algorithm
- $Var(\widehat{m}_{max}) = \frac{1}{2}\left(\frac{\kappa}{s}\right)^2$, where $\kappa = \sqrt{\frac{2}{n}\sum_{i=1}^{n}(q_i)^2}$

▣ On the Use of the CDF-Fitting Procedure

- ❖ These methods assumes that the CDF, $F_M(m)$ is known
- ❖ If so, in spite of efforts up to now, there is no reason to stick to this procedure
- ❖ Instead, we can use the parametric procedures

Iterative Scheme for of β & m_{max}

▣ Inter-Linkage of b & Mmax

❖ In parametric models, they are linked each other

➢ Estimation of b

$$\frac{1}{\hat{\beta}} = \bar{m} - \frac{m_{min} - m_{max}\exp[-\hat{\beta}(m_{max} - m_{min})]}{1 - \exp[-\hat{\beta}(m_{max} - m_{min})]}$$

➢ Estimation of m_{max}

$$\Delta_n = \frac{1 - \exp\left[-\beta\left(m_{max}^{obs} - m_{min}\right)\right]}{n\beta\exp\left[-\beta\left(m_{max}^{obs} - m_{min}\right)\right]}$$

$$\Delta_n = \int_{m_{min}}^{m_{max}} \left[\frac{1 - \exp[-\beta(m - m_{min})]}{1 - \exp[-\beta(m_{max} - m_{min})]}\right]^n$$

❖ To estimate one, the information of the other is necessary

▣ Simultaneous Estimation

❖ Iterative scheme by Noh (2014)

Step 1: estimate β first with observed m_{max}

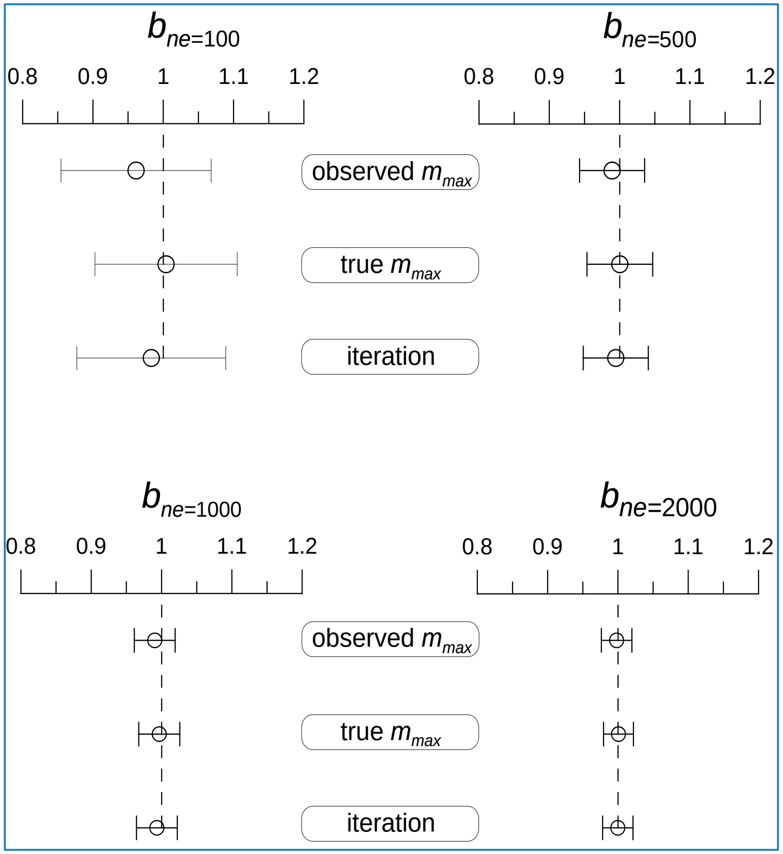
Step 2: estimate m_{max} using β estimated in Step 1

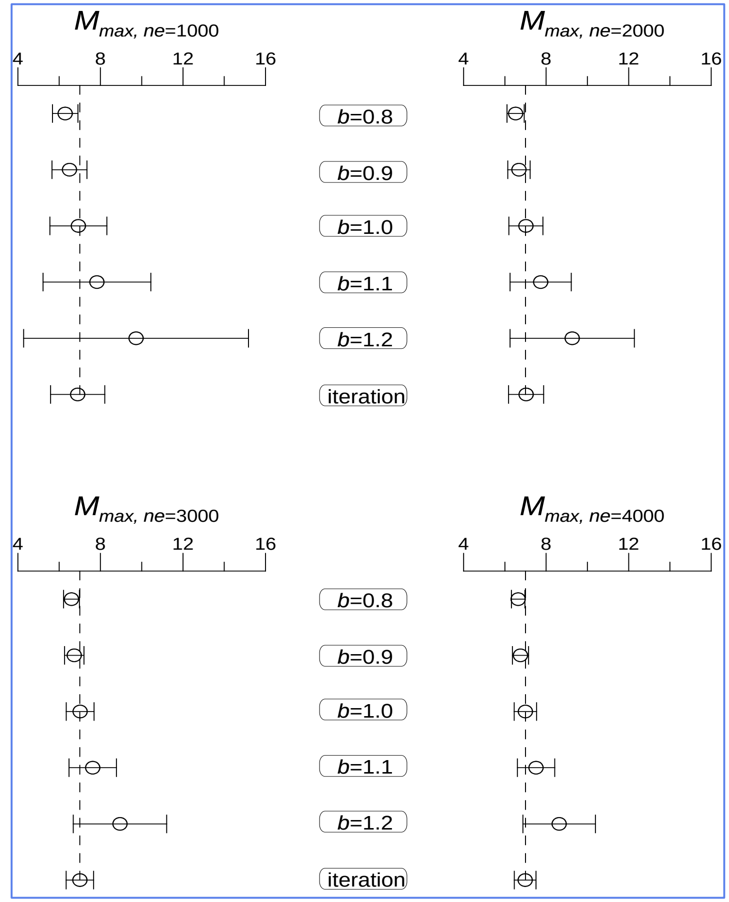
Step 3: re-estimate β using m_{max} estimated in Step 2

Step 4: re-estimate m_{max} using β estimated in Step 3

Step 5: repeat Steps 3 and 4 until certain exit conditions are met

❖ Performance of Iterative Scheme (Noh, 2014)





❖ Recommendations (Noh, 2014)

- Estimate b first,
  - M_{max}^{obs} can effectively replace the unknown M_{max}^{true}
- Then estimate m_{max}
- Ex: Weichert (1980) used m_{max}^{obs} in place of m_{max}

❖ Better Estimate by Iterative scheme

- Starting with b estimation first with m_{max}^{obs} for m_{max}

Chapter 7
Use of Geologic and Geodetic Information

Estimation of Annual Rates

▣ Bridge between Geologic and Seismic Information

- ❖ Seismic moment: $M_0 = \mu A_r \widetilde{D}_r$ (Aki, 1966)
  - ➢ μ: rigidity, ~3x10^{11} dyne/cm^2
  - ➢ A_r: rupture area on a fault plane undergoing slip during an earthquake
  - ➢ $\widetilde{D}_r$: average displacement over the rupture area, i.e.,
    - ▪ $\widetilde{D}_r = \frac{1}{A_r}\int_{A_r} D_r \, dA$, where D_r is a displacement at a rupture point.
  - ➢ $\widetilde{D}_r = M_0/\mu A_r$
- ❖ If little seismic information
  - ➢ $\mu A_r \widetilde{D}_r$ can be used to estimate the amount of seismic moment release
- ❖ If geologic and seismic information available
  - ➢ Estimates are confirmed through comparison

Extension to Whole Fault Surface

❖ Seismic moment rate (Brune, 1968)

- $\widetilde{D}_f = \frac{1}{A_f}\int_{A_f} D_r\, dA = \frac{1}{A_f}\int_{A_r} D_r\, dA = \frac{A_r \widetilde{D}_r}{A_f} = \frac{1}{A_f}\frac{M_0}{\mu} = \frac{M_0}{\mu A_f}$
- Total average slip: $\sum \widetilde{D}_f = \frac{1}{\mu A_f}\sum M_0$
- Total moment rate: $\dot{M}_0^T = \mu A_f S$
  - $\dot{M}_0^T = \frac{1}{T}\sum M_0$: total moment rate during a period T
  - $S = \frac{1}{T}\sum \widetilde{D}_f$: average slip rate over the whole fault plane

Moment Magnitude

❖ $\log M_0 = cm + d$

- c=1.5 & d=16.05 (Hanks and Kanamori, 1979)
- $M_0 = 10^{cm+d} = e^{\gamma m + \delta}$

▣ Slip Rate Constraint

❖ Exponential Distribution

- Gutenberg-Richter relation (Richter, 1958)
  - $\log N(m) = a - bm$ or $N(m) = N^0 e^{-\beta(m-m_0)}$
    - $N^0 = 10^{a-bm_0}$: the number of earthquakes greater than m^0
- Earthquake occurrence density for total N^0 events in $[m^0,\infty)$
  - $n(m) = -\frac{dN(m)}{dm} = N^0 \beta e^{-\beta(m-m^0)}$
- Earthquake occurrence density for total N^0 events in $[m^0,m^u]$
  - Normalization: $k\int_{m^0}^{m^u} n(m)dm = N^0 \rightarrow k\int_{m^0}^{m^u} \beta e^{-\beta(m-m^0)} dm = 1$

    $\therefore k = \left[1 - e^{-\beta(m^u-m^0)}\right]^{-1}$

  - $$n(m) = \begin{cases} \frac{N^0 \beta e^{-\beta(m-m^0)}}{1-e^{-\beta(m^u-m^0)}} & m < m^0 \\ \frac{N^0 \beta e^{-\beta(m-m^0)}}{1-e^{-\beta(m^u-m^0)}}, & m^0 \le m \le m^u \\ 0, & m > m^u \end{cases} \quad (7\text{-}1)$$

Slip Rate Constraint

Exponential Distribution (continued)

- Cumulative number of earthquakes in $[m^0, m^u]$

  - $N(m) = \begin{cases} \frac{N^0\left[e^{-\beta(m-m^0)} - e^{-\beta(m^u-m^0)}\right]}{1-e^{-\beta(m^u-m^0)}} & m < m^0 \\ \frac{N^0\left[e^{-\beta(m-m^0)} - e^{-\beta(m^u-m^0)}\right]}{1-e^{-\beta(m^u-m^0)}}, & m^0 \le m \le m^u \\ 0, & m > m^u \end{cases}$

    e.g., Youngs & Coppersmith (1985)

- Total moment rate during a period T

  - $\dot{M}_0^T = \int_{-\infty}^{\dot{m}} \dot{n}(m) M_0(m) dm$, or (7-2)
  - $\mu A_f S = b\dot{N}^0 M_0^u e^{-\beta(m^u-m^0)} / (c-b)\left(1 - e^{-\beta(m^u-m^0)}\right)$, or
  - $\dot{N}^0 = \frac{\mu A_f S(c-b)\left(1-e^{-\beta(m^u-m^0)}\right)}{b M_0^u e^{-\beta(m^u-m^0)}}$

    where $c > b$ and $M_0^u = M_0(m^u)$ (Youngs & Coppersmith, 1985)

♣ It is worth noting:

- ❖ From (7-1), $n(m)$ can be expressed by PDF: $n(m) = N^0 f_M(m)$
- ❖ But $f_M(m)$ should not be interpreted by a PDF because the integration in (7-2) extends to $-\infty$, below m^0
  - ➢ $f_M(m)$ here is just a function that has the same functional form as the PDF
- ❖ Nevertheless, the analogy to a PDF is quite useful when only the PDF is defined
- ❖ Example: Delta distribution: $f_M(m) = \delta(m - m_p)$
  - ➢ $\dot{M}_0^T = \mu A_f S = \int_{-\infty}^{\infty} \dot{n}(m) M_0(m) dm$

$$= \int_{-\infty}^{\dot{m}} \dot{N}^0 f_M(m) M_0(m) dm$$

$$= \dot{N}^0 \int_{-\infty}^{\infty} \delta(m - m_p) M_0(m) dm$$

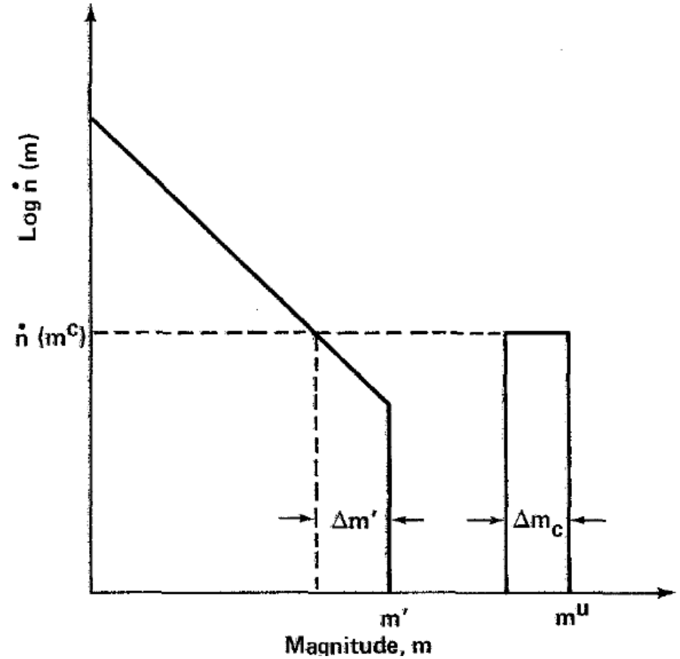
$$= \dot{N}^0 M_0(m_p) \qquad \therefore \dot{N}^0 = \mu A_f S / \, M_0(m_p)$$

- ❖ Conversely, we can find $f_M(m)$ from the formula of $n(m)$

▣ Characteristic Earthquake Model

❖ Schwartz and Coppersmith (1985)

- $\Delta m_c = \frac{1}{2}$
- $m' = m^u - \Delta m_c = m^u - \frac{1}{2}$
- $\Delta m' = 1 \rightarrow \dot{n}^c \equiv \dot{n}(m^c) = \dot{n}(m' - 1)$



❖ Let $N^0 = N^L + N^U$ (7-3)

- N^L : the number of event in $[m^0, m']$
- N^U : the number of event in $[m', m^u]$

❖ From the right figure, we see that

- $N^U = \Delta m_c \, n^c = n^c/2$ (dropping the dot-hat)

❖ Using (7-1) and $n^c = n(m' - 1)$

- $$n(m) = \begin{cases} \dfrac{N^L \beta e^{-\beta(m-m^0)}}{1-e^{-\beta(m'-m^0)}} \\ n^c, \end{cases} = \begin{cases} \dfrac{N^L \beta e^{-\beta(m-m^0)}}{1-e^{-\beta(m'-m^0)}}, & m^0 \le m \le m' \\ \dfrac{N^L \beta e^{-\beta(m'-m^0-1)}}{1-e^{-\beta(m'-m^0)}}, & m' \le m \le m^u \end{cases} \qquad (7\text{-}4)$$

▣ Characteristic Earthquake Model (continued)

❖ Since $n^c = n(m' - 1) = \frac{N^L \beta e^{-\beta(m'-m^0-1)}}{1-e^{-\beta(m'-m^0)}}$

➢ $N^U = \Delta m_c\, n^c = \frac{n^c}{2} = \frac{N^L \beta e^{-\beta(m'-m^0-1)}}{2\left[1-e^{-\beta(m'-m^0)}\right]} = N^L q \qquad \because q \equiv \frac{\beta e^{-\beta(m'-m^0-1)}}{2\left[1-e^{-\beta(m'-m^0)}\right]}$

➢ $N^0 = N^L + N^U = N^L(1+q) \qquad\qquad \therefore N^L = N^0/(1+q)$

❖ Inserting N^L into (7-4)

➢ $$n(m) = \begin{cases} \frac{N^0}{(1+q)} \frac{\beta e^{-\beta(m-m^0)}}{[1-e^{-\beta(m'-m^0)}]} = N^0 k' \beta e^{-\beta(m-m^0)}, & m^0 \le m \le m' \\ \frac{N^0}{(1+q)} \frac{\beta e^{-\beta(m'-m^0-1)}}{[1-e^{-\beta(m'-m^0)}]} = N^0 k' \beta e^{-\beta(m'-m^0-1)}, & m' \le m \le m^u \end{cases}$$

- where $k' = \left[(1+q)\left(1-e^{-\beta(m'-m^0)}\right)\right]^{-1}$

▣ Characteristic Earthquake Model (continued)

❖ Substituting m' by $m^u - 1/2$

➢ $n(m) = \begin{cases} N^0 k' \beta e^{-\beta(m-m^0)}, & m^0 \le m \le m^u - 1/2 \\ N^0 k' \beta e^{-\beta(m^u-m^0-3/2)}, & m^u - 1/2 \le m \le m^u \end{cases}$

- where $q = \frac{\beta e^{-\beta(m^u-m^0-3/2)}}{2\left[1-e^{-\beta(m^u-m^0-1/2)}\right]}$ and $k' = \left[(1+q)\left(1-e^{-\beta(m^u-m^0-1/2)}\right)\right]^{-1}$

❖ As a by-product, we obtained the PDF

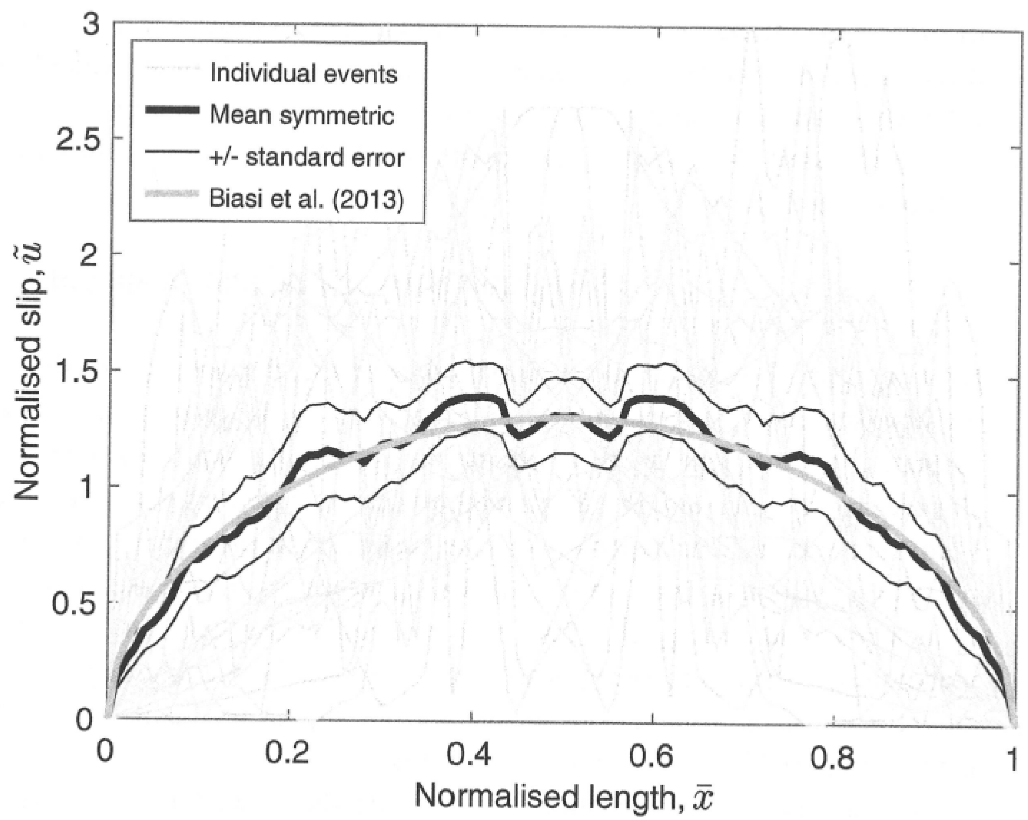
➢ $f_M(m) = \frac{n(m)}{N^0} = \begin{cases} k' \beta e^{-\beta(m-m^0)}, & m^0 \le m \le m^u - 1/2 \\ k' \beta e^{-\beta(m^u-m^0-3/2)}, & m^u - 1/2 \le m \le m^u \end{cases}$

❖ Total moment rate

➢ $\dot{M}_0^T = \int_{-\infty}^{\dot{m}} \dot{n}(m) M_0(m) dm$, or

➢ $\frac{\mu A_f S}{\dot{N}^0} = \frac{k' b M_0(m^u) e^{-\beta(m^u-m^0)}}{c-b} + \sinh(\gamma/4) \frac{2k' b M_0(m^u-1/4) e^{-\beta(m^u-m^0-3/2)}}{c}$

▣ Displacement Distribution on Fault Plane



❖ Normalized slip: $\tilde{u}(\bar{x}) = \frac{u}{\bar{u}} = 1.3\sin^{1/2}(\pi\bar{x})$ (Biasi *et al.*, 2013)

- $\bar{u}$: average slip over the whole fault length
- $\bar{x}$: normalized fault length, $\frac{L}{L_0}$

Estimation of m_{max}

▣ Assumption

- ❖ Growth of the fault dimension due to the occurrence of earthquakes is negligible to small

▣ Use of Geologic and Geodetic Data

- ❖ m_{max} is observed when the whole fault surface is ruptured
- ❖ Empirical relations on the magnitude-rupture length or magnitude-rupture area can be used for the estimation of m_{max}

Chapter 8
Topical Issues

Effect of Catalog Combination

▣ Purpose

- ❖ To increase catalog size for stable estimation of seismicity parameter by extending spatial and/or temporal domains

▣ Case study (Noh, 2020)

- ❖ 3,255 events of M0.1~M5.2 from KMA catalogs for
  - ➢ Period: 1981~2015
  - ➢ Events designated as 'domestic' by KMA
- ❖ Sub-catalogs
  - ➢ Sub-catalog **SL** includes the events occurred in the land of South Korea
  - ➢ Sub-catalog **NL** includes the events occurred in the land of north Korea
  - ➢ Sub-catalog **AO** includes the off-shore events

❖ Estimates of m_c

- Estimates of m_c are high even for the SL, considering the Korean seismic network density
- m_c for the AO and the NL are lager than that for the inland events SL
- m_c for the sub-catalogs (SL+AO) or (SL+AO+NL) is much higher than those for the sub-catalog SL as well as the sub-catalog AO or the sub-catalog NL

Catalog	m_c		m_{max}		b	
	mean	s.d.	mean	s.d.	mean	s.d.
SL	2.8	0.22	5.1	0.55	1.13	0.173
AO	3.2	0.54	5.3	0.14	0.778	0.194
NL	3.1	0.31	4.8	0.32	1.298	0.415
SL+AO	3.6	0.45	5.3	0.15	0.838	0.274
SL+AO+NL	3.8	0.26	5.3	0.19	0.818	0.256

- There exits a trade-off between the completeness and the spatiotemporal coverage of an earthquake catalog
  - To enhance the completeness of an earthquake catalog, divide the catalog into sub-catalogs considering the spatiotemporal detectability of the seismic network
  - Or, one may combine several catalogs to cover a larger region or a longer period at the expense of catalog completeness

Earthquake Double Counting

▣ Types of Seismic Sources

- ❖ Fault source
  - ➢ A fault capable of generating earthquakes
- ❖ Area (Volume) source
  - ➢ A zone where earthquake occurs but the faults responsible those earthquakes are not identified
  - ➢ Could be a large background source, or further divided into several area sources depending on the difference in seismic activities

▣ Spatial Overlapping

- ❖ A fault source generally passes through one or more area sources
- ❖ Those earthquakes counted in for a fault source should not be counted in for the area sources again that contain the fault source
  - ➢ If a new fault source added, the seismicity of all surrounding area sources should be re-assessed

▣ Practical Limits

- ❖ Important seismic parameters to be re-assessed
  - ➢ Annual rate, Richter-b, m_{max}
- ❖ Difficulty in separation of earthquakes
  - ➢ Complete separation of earthquakes of a fault source from the surrounding area sources is impossible due to the uncertainties of the earthquake location and the subsurface structure of fault
  - ➢ Especially, the earthquake location is more uncertain for smaller and older earthquakes
  - ➢ There are some cases where all the large earthquakes, say, larger than M=6.5 are attributed to fault sources
- ❖ Difficulty in the Quaternary faults in Korea
  - ➢ They have been identified solely based on surface geological investigation
  - ➢ There are big uncertainties in the seismic parameter assessed from the geological observation only

▣ Valid Principles

❖ Axiomatic proposition

- There has been a fault. Therefore, finding out the fault does not change the past earthquake history.

$$\sum_{i=1}^{N_b} \nu_i^b = \sum_{j=1}^{N_a} \nu_j^a \qquad (1)$$

  - where N_b and ν_i^b are the number of sources and annual rate of the i-th source before a new fault source is added, and
  - N_a and ν_j^a are the number of sources and annual rate of the j-th source after a new fault source is added

❖ Limit of the axiomatic proposition

- It does not separate earthquakes themselves, but just annul rates
- Thus, it offers no information necessary for re-assessment of the Richter-b and m_{max}

❖ Re-assessment of area sources

- Annual rates
  - Fault: annual rate can be estimated from the geodetic information or paleo-seismic survey
  - Area: annul rates of surrounding area sources can be corrected to the remaining amount of annual rate

▣ Valid Principles

❖ Re-assessment of area sources (continued)

➢ m_{max}

- Fault: m_{max} can be estimated from the geologic information
- Area: m_{max} of an area source is estimated from the earthquake catalog
  - Since the m_{max} estimate is sensitive to the large observed earthquakes, re-assessment of m_{max} of an area source is of particular importance after some large earthquake are assigned to a fault source
  - Re-assessment of m_{max} is possible only when earthquakes themselves were separated

➢ Richter-b

- As long as earthquakes themselves are not separated, the re-assessment of the Richter-b is not possible
- Fortunately, the Richter-b varies little among seismic sources and the separation of earthquakes do not always results in the change of the Richter-b
- It is not so dangerous to use the Richter-b of nearby sources
- This procedure is valid for a fault source as well as an area source

▣ Example Calculation of PSHA (Noh, 2023)

❖ Source map & sites

Identification of fault	Source	$m_{\min}$	$m_{\max}$	$\nu_{m \geq 5}$	Richter−b	Depth	Dip
Before	Area	5.0	7.5	8.0E-2	1.0	5-20 km	-
After	Area	5.0	6.0	3.0E-2	1.0	5-20 km	-
	Fault	5.0	7.5	5.0E-2	1.0	5-20 km	45°SE

❖ GMM: Sadigh et al. (1997), no variability

❖ Spectral frequencies: PGA @ 100 Hz

❖ GM levels: 10 values at

➢ 50, 100, 150, 200, 250, 300, 350, 400, 450, 500 gals

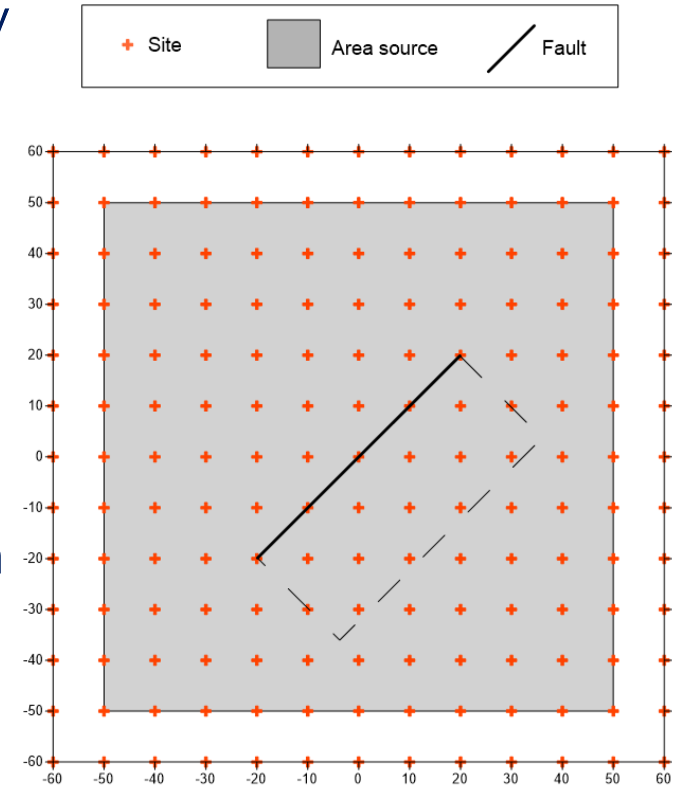
❖ Magnitude-Rupture relation

➢ For length (km): $\log L = \frac{m}{2} - 1.85$

❖ Truncated exponential mag. distribution

❖ Uniform distribution for focal depths

❖ Aspect ratio: 2



Site-source configuration

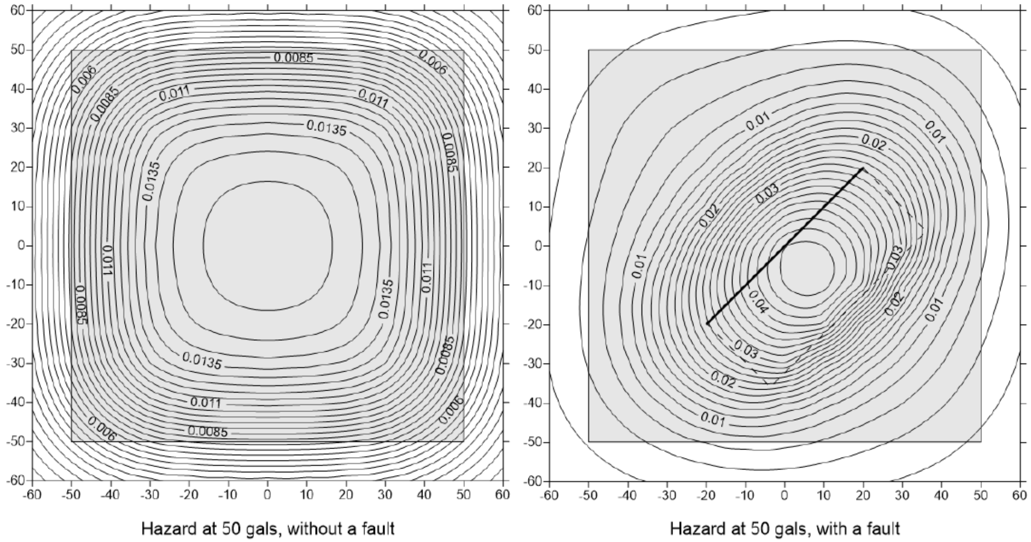


Fig. 2. Spatial distribution of hazard at 50 gals

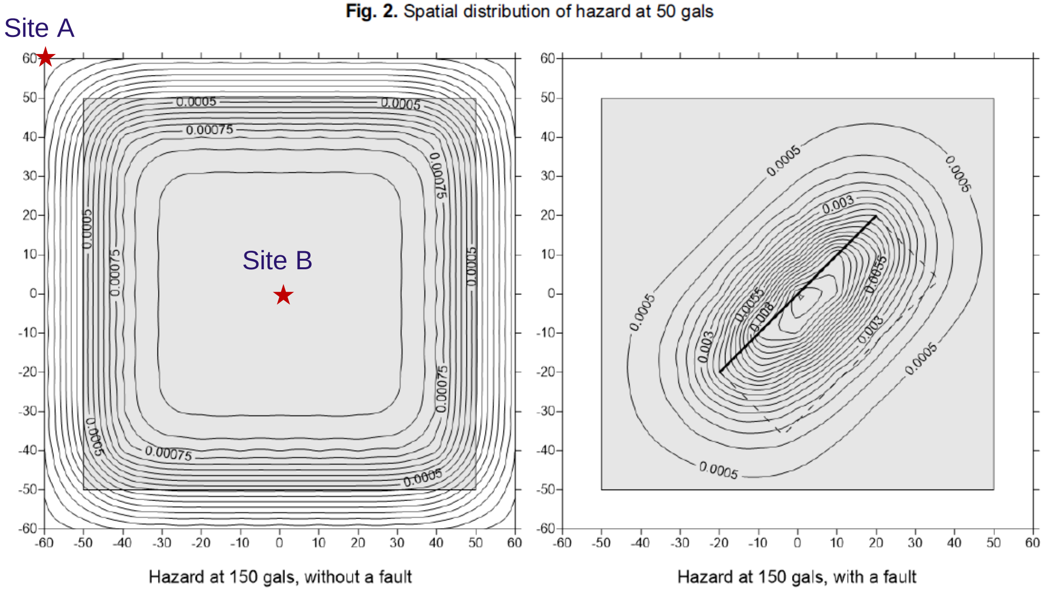
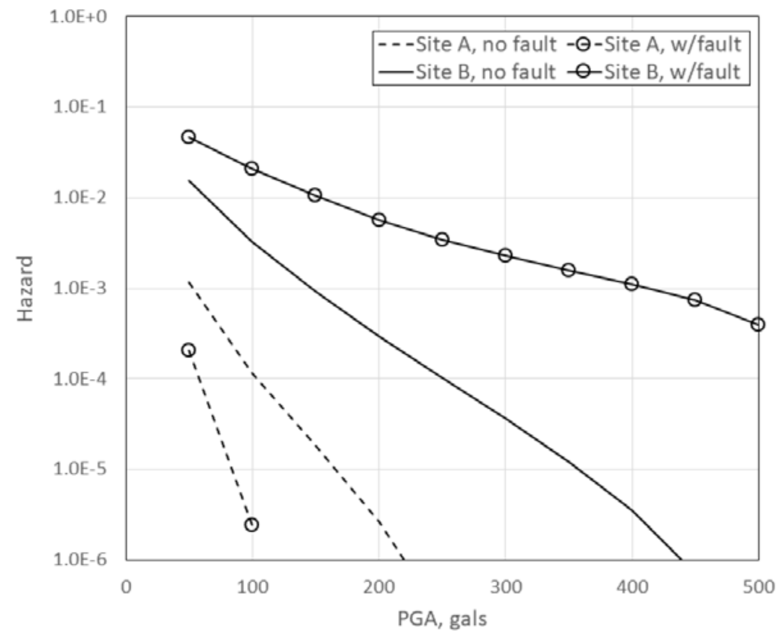


Fig. 3. Spatial distribution of hazard at 150 gals



References

노명현, 2023. 지진의 이중산입에 대한 소고, 한국지진공학회 논문집, 27, 157-162.

노명현, 이상국, 최강룡, 2000. 지진활동 매개변수 추정을 위한 기상청 지진목록의 최소 규모 분석, 지구물리와 물리탐사, 3, 261-268.

Aki, K., 1965. Maximum likelihood estimate of b in the formula $\log N = a - bM$ and its confidence limits, Bull. Earthquake Res. Inst., Tokyo University, 43, 237-239.

Aki, K., 1966. Generation and propagation of G-waves from the Niigata earthquake of June 19, 1964. 2. Estimation of earthquake movement, released energy, and stress-strain drop from G-wave spectrum, Bull. Earthquake Res. Inst., Tokyo Univ. 44, 23-88.

Amorese, D., 2007. Applying a change-point detection method on frequency-magnitude distributions, Bull. Seism. Soc. Am., 97, 1742-1749.

Bender, B., 1983. Maximum likelihood estimation of b values for magnitude grouped data, Bull. Seism. Soc. Am., 73, 831-851.

Berrill, J.B. and R.D. Davis, 1980. Maximum entropy and the magnitude distribution, Bull. Seism. Soc. Am., 70, 1823-1831.

Biasi, G.P., Weldnon, R.J.I., and Dawson, T.E., 2013. Appendix F: Distribution of Slip in Ruptures. Open File Report 2013-1165. USGS.

Brune, J.N., 1968. Seismic moment, seismicity and rate of slip along major fault zones, J. Geophys. Res. 73, 777-784.

Campbell, K.W., 1982. Bayesian analysis of extreme earthquake occurrences. Part I. Probabilistic hazard model, Bull. Seism. Soc. Am., 72, 1869-1705.

Cao, A.M. and S.S. Gao, 2002. Temporal variation of seismic b-values beneath northeastern Japan island arc, Geophys. Res. Lett., 29, doi 10.1029/2001GL013775.

Cooke, P. (1979), Statistical inference for bounds of random variables, *Biometrika,* 66, 2, 367-374, DOI: 10.1093/biomet/66.2.367.

Cornell, C.A., 1972. Bayesian statistical decision theory and reliability-based design, Proceedings of the International Conference on Structural Safety and Reliability, April 9-11, 1969, Washington, D.C., Smithsonian Institute, 47-66.

Cornell, C.A. and E.H. Van Marke, 1969. The major influences on seismic risk, Proceedings Third World Conference on Earthquake Engineering, Santiago, Chile, A-1, 69-93.

Cramér, H., 1961. Mathematical Methods of Statistics, 2nd ed., Princeton University Press, Princeton.

Gnedenko, B. (1943), Sur la distribution limite du terme maximum d'une série aléatoire, *Ann. Math*. 44, 3, 423-453 (in French).

Hanks, T.C. and H. Kanamori, 1979. A moment magnitude scale, J. Geophys. Res., 84, 2348-2350.

Johnston, A.C., Coppersmith, K.J., Kanter, L.R., and Cornell, C.A., 1994. The earthquakes of stable continental regions: final report submitted to Electric Power Research Institute: TR-102261.

Karnik, V.K., 1971. Seismicity of the European area, Part 2. Academia, Publishing House of the Czechoslovak Academy of Sciences, Praha, Czechoslovakia, 123-169.

Kijko, A., 2004. Estimation of the maximum earthquake magnitude, m_{max}. Pure and Applied Geophysics, 161, 1-27.

Kijko, A., and M.A. Sellevoll (1989), Estimation of earthquake hazard parameters from incomplete data files. Part I. Utilization of extreme and complete catalogs with different threshold magnitudes, Bull. Seism. Soc. Am. 79, 3, 645-654.

Kijko, A. and A. Smit, 2012. Extension of the Aki-Utsu b-value for incomplete catalogs, Bull. Seism. Soc. Am., 102, 1283-1287.

Noh, M., 2014. A parametric estimation of Richter-b and m_{max} from an earthquake catalog, Geosciences Jour., 18, 339-345.

Noh, M., 2016. On the Poisson process of the Korean earthquakes, Geosciences Jour., 20, 775-779.

Noh, M., 2019. Assessment of the completeness of earthquake catalogs, Geosciences Jour., 23, 253-263.

Noh, M., 2020. Effect of combining catalogs with different completeness: EGU2020-1749, EGU General Assembly 2020, Vienna.

Ogata, Y. and K. Katsura, 1993. Analysis of temporal and spatial heterogeneity of magnitude frequency distribution inferred from earthquake catalogs, Geophys. J. Int., 113, 727-738.

Page, R., 1968. Aftershocks and microaftershocks of the great Alaska earthquake of 1964, Bull. Seism. Soc. Am., 58, 1131-1168.

Pisarenko, V.F., Lyubushin, A.A., Lysenko, V.B., and Golubieav, T.V., 1996. Statistical estimation of seismic hazard parameters: maximum possible magnitude and related parameters, Bull. Seism. Soc. Am., 86, 691-700.

Press, W.H., B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling (1994) *Numerical Recipes: The Art of Scientific Programming*, Cambridge University Press, New York.

Quenouille, M.H. (1956), Notes on bias in estimation, *Biometrika,* 43, 3-4, 353-360, DOI: 10.1093/biomet/43.3-4.353.

Richter, C.F., 1975. Elementary Seismology, W.H. Freeman and Company, San Francisco, California.

Robson, D.S., and J.H. Whitlock (1964), Estimation of a truncation point, *Biometrika,* 51, 1-2, 33-39, DOI: 10.1093/biomet/51.1-2.33.

Rydelek, P.A. and Sacks, I.S., 1989. Testing the completeness of earthquake catalogs and the hypothesis of self-similarity, Nature, 337, 251-253.

Sadigh, K., Chang C.-Y., Egan J.A., Makdisi F., and Youngs R.R., 1997. Attenuation relationships for shallow crustal earthquakes based on California Strong Motion Data, Bull. Seism. Soc. Am., 68, 180-189.

Schwartz D.P. and Coppersmith, K.J., 1984. Fault behavior and characteristic earthquakes: examples from the Wasatch and San Andreas faults, J. Geophys. Res., 89, 5681-5698.

Tate, R.F., 1959. Unbiased Estimation: Function of Location and Scale Parameters, Ann. Math. Statist., 30, 331–366.

Utsu, T., 1965. A method for determining the value of b in formula $\log N = a - bM$ showing the magnitude-frequency relation for earthquakes, Geophys. Bull. Hokkaido University, 13, 99-103.

Utsu, T., 1966. A statistical significance test of the difference in b-value between two earthquake groups, J. Phys. Earth, 14, 37-40.

Weichert, D.H., 1980. Estimation of the earthquake recurrence parameters for unequal observation periods for different magnitudes, Bull. Seism. Soc. Am., 70, 1337-1346.

Wiemer, S. and M. Wyss, 2000. Minimum magnitude of complete reporting in earthquake catalogs: examples from Alaska, the Wester United States, and Japan, Bull. Seism. Soc. Am., 84, 974-1002

Wells, D.H. and Coppersmith, K.J., 1994. New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement, Bull. Seism. Soc., Am., 84, 974-1002.

Woessmer, J. and S. Wiemer, 2005. Assessing the quality of earthquake catalogues: estimating the magnitude of completeness and its uncertainty, Bull. Seism. Soc. Am., 95, 684-698.

Youngs, R.R. and Coppersmith, K.J., 1985. Implications of fault slip rates and earthquake recurrence models to probabilistic seismic hazard estimates, Bull. Seism. Soc. Am., 75(4), 939-964.