Mahler's Guide to Statistics

Information in bold or sections whose title is in bold are more important for passing the exam. Information presented in italics (and sections whose titles are in italics) is much less likely to be needed to directly answer exam questions, and should be skipped on first reading.

Solutions are posted at http://www.neas-seminars.com.

Table of Contents continued on the next page.

Highly Recommended problems are double underlined.
Recommended problems are underlined.[1]

"Mahler's Guide to Statistics" covers CAS 3L Learning Objective: D.

[1] Solutions are posted at http://www.neas-seminars.com
Note that problems include both some written by me and some from past exams.

CAS Exam 3 Questions by Section of this Study Aid[2]

	3	3	3	3	3	3	3L	3L	3L	3L
Sec.	5/05	11/05	5/06	11/06	5/07	11/07	5/08	11/08	5/09	11/09
1										
2	19		1	3		5			17	17
3										
4	18, 20	1, 4	2	2	10, 11	6	3	4, 6	19	18, 19
5										
6							1		18	
7										
8										
9		5				10				20
10		8								
11					18	9	4			
12				5	19			9		23
13										
14										
15				7			6		22	
16					29			11		
17										
18						7				
19	24	3, 7	6	6	23		5		20	
20	22, 23				22					21
21										
22			7		30	11			21	25
23										
24			5	4		8		8		
25								7		22
26										
27	27		9	8			9		25	
28		9								
29										
30					31					
31										
32										
33										
34	25	2	8	9		12	7, 8	10	23	24
35										
36	21	6	3, 4	1		4	2	5	24	
37										
38										

[2] Statistics was added to the syllabus in 2005. Statistics was also on the syllabus prior to 2000.

Section 1, Introduction

This Statistics material was added to the CAS 3 Exam with the Spring 2005 exam.
Prior to 2000, this Statistics material was on the joint CAS Part 2 / SOA Course 110.[3]
While questions from these old exams give us some idea of the type of question you will be asked, there have been some changes in style of questions asked.

There is not always a clear boundary between this Statistics material and the Probability material on joint Exam 1/P.
If you have studied Probability, it would not hurt to briefly review that material.

There is not always a clear boundary between this Statistics material and the Regression material which is now part of Validation by Education Experience. The questions on this exam should be limited to least squares fitting of straight lines and other simple ideas.

There is not always a clear boundary between this Statistics material and the material on fitting models on joint Exam 4/C. On fitting via method of moments and maximum likelihood, the questions on this exam should be limited to more basic applications of those ideas.
With respect to the testing of fits via the Chi-Square Goodness of Fit Test and the Likelihood Ratio Test, any questions on this exam should be limited to very basic applications of those ideas.

A **statistic** is a function of observations of either a sample from a single random variable or samples from random variables. Examples of statistics include: the sample mean, the sample variance, the sample third moment, the minimum, the maximum, and the sample median.

A statistic will have a distribution of possible values with corresponding probabilities.
If one can determine the distribution or approximate distribution of a given statistic under certain assumptions, then one can perform tests of hypotheses. Among the distributions that will be used for this purpose are: the Normal, t-distribution, Chi-Square, and F-Distribution. Tables of each of these distributions should be attached to the exam.

The CAS Syllabus provides some guidance on the expected split of questions by topic within Statistics. Also there are several recent exams that one can review. However, you should still use my designation of certain Statistics material as more important with appropriate caution.

[3] Relevant questions from these old exams are included throughout this study guide.

Since there is no single designated textbook to use, the boundaries of the syllabus are not as clear as they might be. Therefore, I have included a significant number of italicized sections. Be sure not to get bogged down; concentrate on important sections your first time through.

CAS Learning Objectives for Statistics (33% to 37% of the exam):

Learning Objective #1:
Perform point estimation of statistical parameters using the following statistical methods:

- Maximum likelihood estimation ("MLE")
- Method of moments

Apply criteria to the estimates such as:

- Consistency
- Unbiasedness
- Minimum variance
- Mean square error

Range of weight: 10-15 percent
Sections of Study Guide: 2, 3, 4, 36
Knowledge Statements:

a. Equations for MLE of mean, variance from a sample
b. Estimation of mean and variance based on sample
c. General equations for MLE of parameters
d. Equations for estimation of parameters using method of moments for means, variances, and higher moments
e. Recognition of consistency property of estimators and alternative measures of consistency
f. Application of criteria for measurement when estimating parameters through minimization of variance, mean square error
g. Definition of statistical bias and recognition of estimators that are unbiased or biased

Learning Objective #2:
Test statistical hypotheses including Type I and Type II errors using:

- Neyman-Pearson lemma
- Likelihood ratio tests

Apply Neyman-Pearson lemma to construct likelihood ratio equation.
Range of weight: 10-15 percent
Sections of Study Guide: 18-20, 24-25

Knowledge Statements:

a. Presentation of fundamental inequalities based on general assumptions and normal assumptions
b. Definition of Type I and Type II errors
c. Significance levels
d. One-sided versus two-sided tests
e. Estimation of sample sizes under normality to control for Type I and Type II errors
f. Determination of critical regions
g. Definition and measurement of likelihood ratio tests
h. Determining parameters and testing using tabular values
i. Recognizing when to apply likelihood ratio tests versus chi-square or other goodness of fit tests (statistics)

Learning Objective #3:

Calculate order statistics of a sample and use critical values from a sampling distributions to test means and variances.

Range of weight: 3-7 percent

Sections of Study Guide: 6-15, 21-22, 34

Knowledge Statements:

a. General form for distribution of nth largest element of a set
b. Application to a given distributional form
c. Recognition of random variables from sample that behave as t-stat or F-stat
d. Determination of parameters when applying these tests and obtaining tabular values
e. Presentation of hypotheses testing from above (t-test and F-test) for mean and variances

Learning Objective #4:

Perform a linear regression using the least squares method.

Range of weight: 3-7 percent

Sections of Study Guide: 26-33

Knowledge Statements:

a. Presentation and calculation of equations for regression statistics

One could outline the Statistics Material as follows:

Fitting
- Method of Moments
- Maximum Likelihood
- Regression

Statistical Tests
- Hypothesis Testing
- Normal Distribution
- Chi-Square Distribution
- t Distribution
- F Distribution

Order Statistics

Properties of Estimators

Tables to be attached to the CAS Exam 3L:

For CAS Exam 3L, a Normal table, the Illustrative Life Table, a Chi-Square table, t-table, and F-table will be attached.

Appendix A and B from Loss Models, giving useful information on loss and frequency distributions, will also attached.

Below is some information on loss and frequency distributions, parameterized as per Loss Models.[4]

Check the CAS webpage for the tables to be attached to your exam.

[4] For Spring 2007, the material on frequency and loss distributions was moved to Exam 4/C.

<u>Binomial Distribution</u>:

Support: x = 0, 1, 2, 3, ..., **m**. Parameters: 1 > q > 0, m ≥ 1. m integer

D. f. : *$F(x) = 1 - \beta(x+1, m-x ; q) = \beta(m-x , x+1 ; 1-q)$* *Incomplete Beta Function*

P. d. f. : $$\mathbf{f(x)} = \frac{m!\ q^x\ (1-q)^{m-x}}{x!\ (m-x)!} = \binom{m}{x} q^x\ (1-q)^{m-x}.$$

Mean = mq

Variance = mq(1-q) **Variance / Mean = 1 - q < 1.**

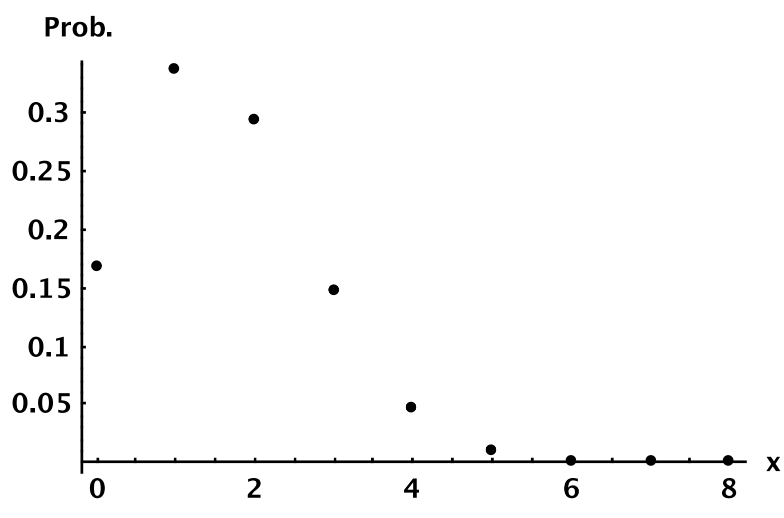
Mode = largest integer in mq + q (if mq + q is an integer, then f(mq + q) = f(mq + q- 1) and both mq + q and mq + q - 1 are modes.)

A Binomial Distribution with m = 1 is a Bernoulli Distribution.

The sum of m independent Bernoulli Distributions with the same q is a Binomial with parameters m and q.

The sum of two independent Binomials with parameters (m_1 , q) and (m_2, q) is also Binomial with parameters $m_1 + m_2$ and q.

Binomial Distribution with m = 8 and q = 0.2:



<u>Using the Functions of the Calculator to Compute Binomial Coefficients</u>:

Using the TI-30X-IIS, the binomial coefficient $\binom{n}{i}$ can be calculated as follows:

n
PRB
▸
nCr
Enter
i
Enter

For example, in order to calculate $\binom{10}{3} = \frac{10!}{3!\ 7!} = 120$:

10
PRB
▸
nCr
Enter
3
Enter

Poisson Distribution:

Support: x = 0, 1, 2, 3, ... Parameters: $\lambda > 0$

D. f. : *$F(x) = 1 - \Gamma(x+1 ; \lambda)$* *Incomplete Gamma Function*

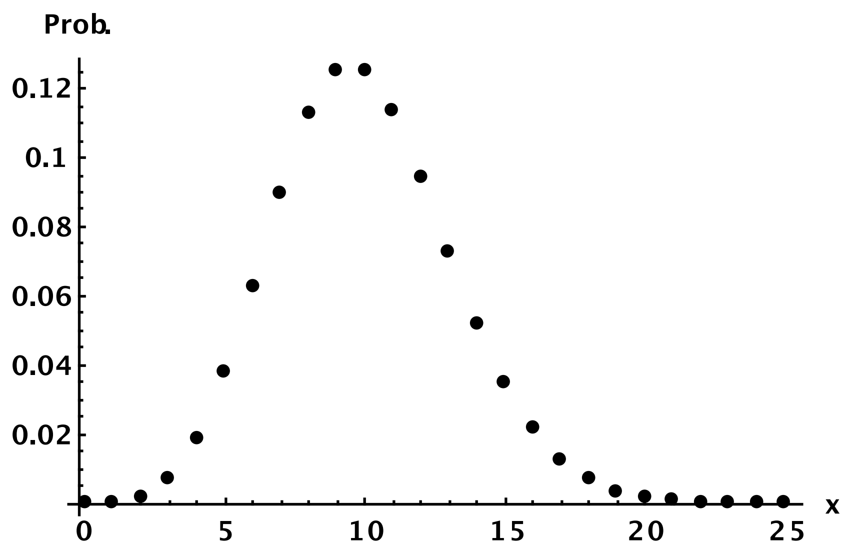
P. d. f. : $\mathbf{f(x) = \lambda^x e^{-\lambda} / x!}$

Mean = λ **Variance = λ** **Variance / Mean = 1.**

Mode = largest integer in λ (if λ is an integer then both λ and λ-1 are modes.)

The sum of two independent variables each of which is Poisson with parameters λ_1 and λ_2 is also Poisson, with parameter $\lambda_1 + \lambda_2$.

A Poisson Distribution for λ = 10:



Geometric Distribution

Support: x = 0, 1, 2, 3, ... Parameters: $\beta > 0$.

D. f. : $F(x) = 1 - \left(\frac{\beta}{1+\beta}\right)^{x+1}$

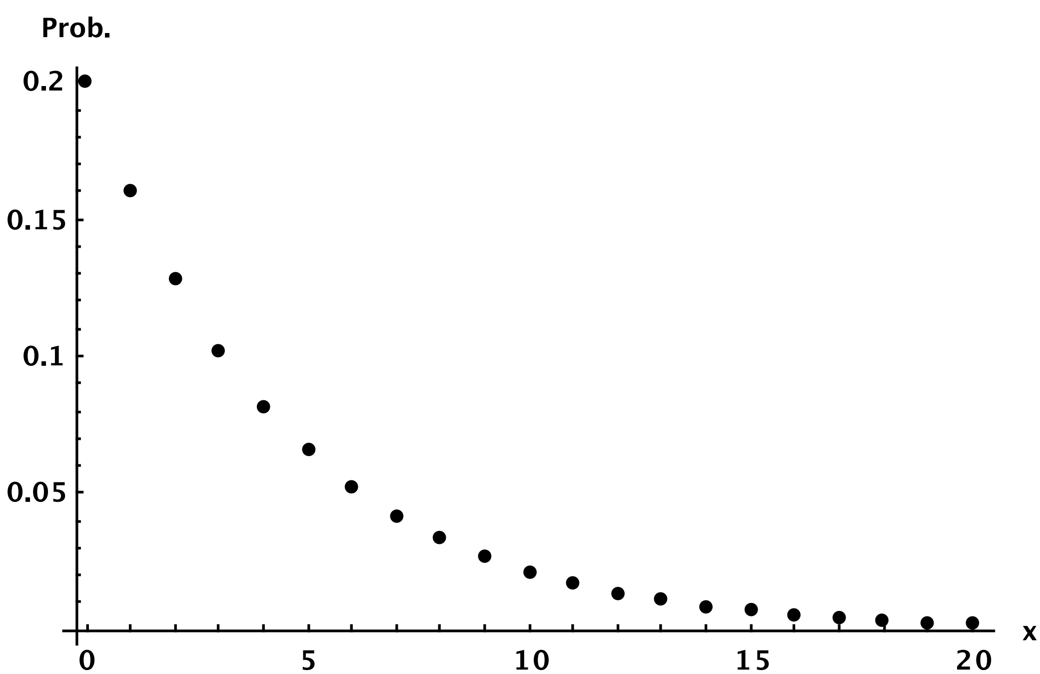
P. d. f. : $\mathbf{f(x)} = \frac{\beta^x}{(1+\beta)^{x+1}}$

f(0) = 1/ (1+β). $f(1) = \beta / (1 + \beta)^2$. $f(2) = \beta^2 / (1 + \beta)^3$. $f(3) = \beta^3 / (1 + \beta)^4$.

Mean = β **Variance = $\beta(1+\beta)$** **Variance / Mean = 1 + β > 1.**

Mode = 0.

A Geometric Distribution for $\beta = 4$:



<u>Negative Binomial Distribution</u>:

Support: $x = 0, 1, 2, 3, \ldots$ Parameters: $\beta > 0$, $r \geq 0$. **r = 1 is a Geometric Distribution**

D. f. : *F(x) = β(r, x+1 ; 1/(1+β)) =1- β(x+1, r ; β/(1+β))* *Incomplete Beta Function*

P. d. f. : $$\mathbf{f(x)} = \frac{r(r+1)\ldots(r+x-1)}{x!}\frac{\beta^x}{(1+\beta)^{x+r}} = \binom{x+r-1}{x}\frac{\beta^x}{(1+\beta)^{x+r}}.$$

$$\mathbf{f(0) = 1/(1+\beta)^r}.$$

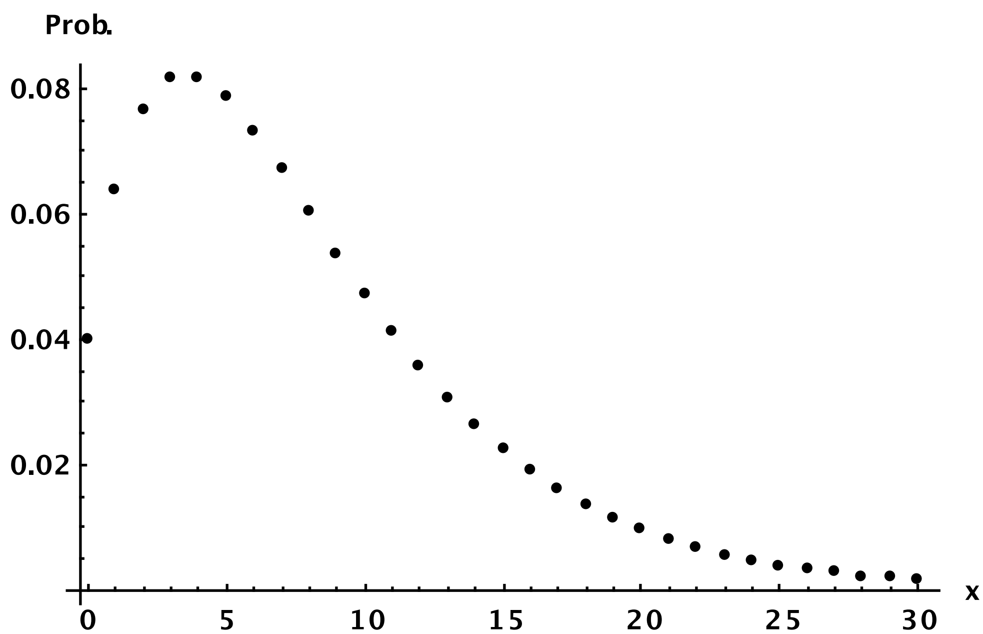
$f(1) = r\beta / (1+\beta)^{r+1}$. $f(2) = \{r(r+1)/2\}\beta^2 / (1+\beta)^{r+2}$. $f(3) = \{r(r+1)(r+2)/6\}\beta^3 / (1+\beta)^{r+3}$.

Mean = rβ **Variance = rβ(1+β)** **Variance / Mean = 1 + β > 1.**

Mode = largest integer in (r-1)β (if (r-1)β is an integer, then both (r-1)β and (r-1)β - 1 are modes.)

The sum of r independent Geometric Distributions with the same β, is a Negative Binomial Distribution with parameters r and β.
X is Negative Binomial(r_1, β) and Y is Negative Binomial(r_2, β), X and Y independent, then X + Y is Negative Binomial($r_1 + r_2$, β).

A Negative Binomial Distribution with $r = 2$ and $\beta = 4$:



Exponential Distribution:

$F(x) = 1 - e^{-x/\theta}$ **$f(x) = e^{-x/\theta} / \theta$, $x > 0$.**

Mean = θ **Variance = θ^2**

Second Moment = $2\theta^2$ $E[X^n] = n!\, \theta^n$

When an Exponential Distribution is truncated and shifted from below, in other words when looks at the nonzero payments excess of a deductible, one gets the same Exponential Distribution, due to its memoryless property.

Single Parameter Pareto Distribution:

$F(x) = 1 - \left(\frac{\theta}{x}\right)^{\alpha}$, $x > \theta$. $f(x) = \frac{\alpha\, \theta^{\alpha}}{x^{\alpha+1}}$, $x > \theta$.

Mean = $\frac{\alpha\, \theta}{\alpha - 1}$, $\alpha > 1$ Variance = $\frac{\alpha\, \theta^2}{(\alpha-1)^2\, (\alpha-2)}$, $\alpha > 2$

$E[X^2]$ = $\frac{\alpha\, \theta^2}{(\alpha-2)}$, $\alpha > 2$ $E[X^n]$ = $\frac{\alpha\, \theta^n}{\alpha - n}$, $\alpha > n$

Pareto Distribution:

$F(x) = 1 - \left(\frac{\theta}{\theta + x}\right)^{\alpha}$, $x > 0$ $f(x) = \frac{\alpha\, \theta^{\alpha}}{(\theta + x)^{\alpha+1}}$, $x > 0$.

Mean = $\frac{\theta}{\alpha - 1}$, $\alpha > 1$. Variance = $\frac{\alpha\, \theta^2}{(\alpha-1)^2\, (\alpha-2)}$, $\alpha > 2$.

$E[X^2]$ = $\frac{2\, \theta^2}{(\alpha-1)\, (\alpha-2)}\}$, $\alpha > 2$. $E[X^n]$ = $\frac{n!\, \theta^n}{\prod_{i=1}^{n} (\alpha - i)} = \frac{n!\, \theta^n}{(\alpha-1)...(\alpha-n)}$, $\alpha > n$

Gamma Distribution:

$F(x) = \Gamma[\alpha ; x/\theta]$ $f(x) = \frac{(x/\theta)^{\alpha} e^{-x/\theta}}{x\,\Gamma(\alpha)} = \frac{x^{\alpha-1} e^{-x/\theta}}{\theta^{\alpha}\,\Gamma(\alpha)}$, $x > 0$.

Mean = $\alpha\theta$ **Variance =** $\alpha\theta^2$ $E[X^2]$ **=** $\theta^2 \alpha(\alpha + 1)$

$E[X^n]$ **=** $\theta^n \prod_{i=0}^{n-1} (\alpha + i) = \theta^n (\alpha)...(\alpha + n - 1) = \theta^n \frac{\Gamma[\alpha + n]}{\Gamma[\alpha]}$.

The sum of n independent identically distributed variables which are Gamma with parameters α and θ is a Gamma distribution with parameters $n\alpha$ and θ.

For α = a positive integer, the Gamma distribution is the sum of α independent variables each of which follows an Exponential distribution.
For $\alpha = 1$ you get the Exponential.

For very large α, the Gamma distribution approaches a symmetric Normal Distribution.

LogNormal Distribution:

If ln(x) follows a Normal Distribution, then x itself follows a LogNormal Distribution.

$F(x) = \Phi\left[\frac{\ln(x) - \mu}{\sigma}\right]$ $f(x) = \frac{\exp\left[-\frac{(\ln(x) - \mu)^2}{2\sigma^2}\right]}{x\,\sigma\,\sqrt{2\pi}}$, $x > 0$.

Mean = $\exp(\mu + 0.5\,\sigma^2)$ Variance = $\exp(2\mu + \sigma^2)\,\{\exp(\sigma^2) - 1\}$

$E[X^2]$ **=** $\exp[2\mu + 2\,\sigma^2]$. $E[X^n]$ **=** $\exp[n\mu + 0.5\,n^2\,\sigma^2]$.

Weibull Distribution:

$F(x) = 1 - \exp\left[-\left(\frac{x}{\theta}\right)^{\tau}\right]$ $f(x) = \frac{\tau \left(\frac{x}{\theta}\right)^{\tau} \exp\left[-\left(\frac{x}{\theta}\right)^{\tau}\right]}{x} = \frac{\tau\,x^{\tau-1} \exp\left[-\left(\frac{x}{\theta}\right)^{\tau}\right]}{\theta^{\tau}}$, $x > 0$.

For $\tau = 1$ you get the Exponential Distribution.

Section 2, Method of Moments

One can fit a distribution to data via the **Method of Moments**, by finding that set of parameters such that the moments match the observed moments.

If one has a single parameter, such as in the case of the Poisson Distribution, then one matches the observed mean to the theoretical mean of the distribution.
In the case of two parameters, one matches the first two moments, or equivalently one matches the mean and variance.

Poisson Distribution:

Fitting via the method of moments is easy for the Poisson, one merely **sets the single parameter λ equal to the observed mean**, $\hat{\lambda} = \bar{X}$.

Exercise: Assume one has observed insureds and gotten the following distribution of insureds by number of claims:

Number of Claims	0	1	2	3	4	5	6	7	8	All
Number of Insureds	17649	4829	1106	229	44	9	4	1	1	23872

Fit this data to a Poisson Distribution via the Method of Moments.
[Solution: By taking the average value of the number of claims, one can calculate that the first moment is: $\bar{X}$ = {(0)(17649) + (1)(4829) + (2)(1106) + (3)(229) + (4)(44) + (5)(9) + (6)(4) + (7)(1) + (8)(1)} / 23872 = 0.3346. Thus the fitted Poisson has λ = 0.3346.]

As discussed in a subsequent section, **for the Poisson the Method of Maximum Likelihood equals the Method of Moments.**

Binomial Distribution, m fixed:

If as is common when fitting a Binomial, m is taken as fixed, then there is only one parameter q, and solving via the method of moments for the Binomial is easy: $\hat{q} = \bar{X} / m$.

As discussed in a subsequent section, for m fixed, this is also the Method of Maximum Likelihood solution for q.

Exercise: Assume one has observed insureds and got the following distribution of insureds by number of claims:

Number of Claims	0	1	2	3	4	5	6	7 & +	All
Number of Insureds	208	357	274	126	31	3	1	0	1000

Fit this data to a Binomial Distribution with m = 6 via the Method of Moments.

[Solution: By taking the average value of the number of claims, one can calculate that the first moment is: $\bar{X}$ = 1428/1000 = 1.428. Then the fitted q = 1.428/6 = .238.

Comment: If this data is from a Binomial Distribution, then we know that m ≥ 6, since for the Binomial Distribution x ≤ m.]

Negative Binomial, r fixed:

If one takes r as fixed in the Negative Binomial, then solving via the method of moments is straightforward: $\bar{X} = r\beta \Rightarrow \hat{\beta} = \bar{X} / r$.

Exercise: For r = 1.5, fit the following data to a Negative Binomial Distribution via the methods of moments.

Number of Claims	0	1	2	3	4	5	6	7	8	All
Number of Insureds	17649	4829	1106	229	44	9	4	1	1	23872

[Solution: $\bar{X}$ = .3346. The fitted $\beta = \bar{X}/r$ = .3346 / 1.5 = .223.]

As discussed in a subsequent section, for r fixed, this is also the Method of Maximum Likelihood solution for β.

Negative Binomial:

Assume one has observed insureds and gotten the following distribution of insureds by number of claims:

Number of Claims	0	1	2	3	4	5	6	7	8	All
Number of Insureds	17649	4829	1106	229	44	9	4	1	1	23872

By taking the average value of the number of claims, one can calculate that the first moment is .3346. By taking the average value of the square of the number of claims observed for each insured, one can calculate that the second moment is: $\{(0^2)(17649) + (1^2)(4829) + (2^2)(1106) + (3^2)(229) + (4^2)(44) + (5^2)(9) + (6^2)(4) + (7^2)(1) + (8^2)(1)\} / 23872 = 0.5236$.

Thus the estimated variance is: $.5236 - .3346^2 = 0.4116$.

Since the estimated variance is (significantly) greater than the estimated mean, it might make sense to fit a Negative Binomial Distribution to this data. Using the method of moments one would try to match the first two moments by fitting the two parameters of the Negative Binomial Distribution r and β.

Exercise: Fit the above data to a Negative Binomial Distribution via the Method of Moments.
[Solution: One can write down two equations in two unknowns, by matching the mean and the variance: $\bar{X} = r\beta$. Variance $= r\beta(1+\beta)$. $\Rightarrow 1 + \beta = \text{Variance}/\bar{X} = .4116/.3346 = 1.230$.
$\Rightarrow \beta = .230$. $\Rightarrow r = \bar{X}/\beta = 0.3346 / .230 = 1.455$.
Comment: One could get the same result by instead matching the means and second moments.
$r\beta = .3346$, and $r\beta(1+\beta) + (r\beta)^2 = .5236$.]

In general for the Negative Binomial, the method of moments consists of writing two equations for the two parameters r and β by matching the first two moments, or equivalently, one can match the mean and variance: $\bar{X} = r\beta$. $E[X^2] - \bar{X}^2 = r\beta(1+\beta)$.

The solution is: $\hat{\beta} = \dfrac{E[X^2] - \bar{X}^2 - \bar{X}}{\bar{X}}$. $\hat{r} = \dfrac{\bar{X}^2}{E[X^2] - \bar{X}^2 - \bar{X}} = \bar{X}/\hat{\beta}$.

One can compare the Negative Binomial distribution fitted via method of moments with $\beta = .230$ and r = 1.455, to the observed distribution:[5]

Number of Claims	Observed	Method of Moments Negative Binomial
0	17,649	17,663.5
1	4,829	4,805.8
2	1,106	1,103.1
3	229	237.6
4	44	49.5
5	9	10.1
6	4	2.0
7	1	0.4
8	1	0.1
9	0	0.0
10	0	0.0
Sum	23872.0	23872.0

The fit via method of moments seems to be a reasonable first approximation.

[5] The densities of the fitted Negative Binomial have been multiplied by the total number of insureds observed.

Exponential Distribution:

Since it has a single parameter, in the case of the Exponential Distribution, one matches the observed mean to the theoretical mean of the distribution.

Exercise: A data set of 10 losses is: 2, 7, 11, 27, 29, 47, 52, 75, 110, 260.
Fit an Exponential Distribution to this data set using the Method of Moments.
[Solution: $\overline{X}$ = (2 + 7 + 11 + 27 + 29 + 47 + 52 + 75 + 110 + 260)/ 10 = 62. θ = 62.]

Two Parameter Distributions:

If one has a distribution with two parameters, then one can either match the first two moments or match the mean and the variance, whichever is easier.[6]

The above data set has a sample second moment of:
$E[X^2] = (2^2 + 7^2 + 11^2 + 27^2 + 29^2 + 47^2 + 52^2 + 75^2 + 110^2 + 260^2)/ 10 = 9198.2$.

In order to fit a Gamma distribution to this data we match the first two moments, since the Gamma has two parameters alpha and theta:
$62 = \alpha\theta$.
$9198.2 = \alpha\ (\alpha+1)\ \theta^2$.
Dividing the second equation by the square of the first equation:
$(\alpha+1)/\alpha = 9198.2/62^2 = 2.393$.
$(\alpha+1)2.393 = \alpha$. Thus $\alpha = 0.7179$.
Therefore $\theta = 86.37$.

Exercise: Fit the Pareto Distribution to this same data.
[Solution: Set the first moments equal: $62 = \frac{\theta}{\alpha - 1}$,

and set the second moments equal: $9198.2 = \frac{2\ \theta^2}{(\alpha - 1)\ (\alpha - 2)}$.

Dividing the second equation by the square of the first, eliminates θ:
$2.393 = 2(\alpha-1)/(\alpha-2)$. $\Rightarrow \alpha = 7.089$. $\Rightarrow \theta = 377.5$.]

[6] One gets the same answer either way.

Exercise: Fit the LogNormal Distribution to this same data.
[Solution: Set the first moments equal: $62 = \exp(\mu + \sigma^2/2)$,
and set the second moments equal: $9198.2 = \exp(2\mu + 2\sigma^2)$.
Dividing the second equation by the square of the first, eliminates mu:
$2.393 = \exp(\sigma^2)$. $\Rightarrow \sigma = 0.934$.
$\mu = \ln(62) - (0.934^2)/2 = 3.691$.]

It is relatively easy to check the results of fitting via the method of moments. For example, for a LogNormal with parameters $\mu = 3.691$ and $\sigma = 0.934$, the first moment is:
$\exp(3.691 + 0.934^2/2) = 62$ and the second moment is: $\exp[2(3.691) + 2(0.934^2)] = 9198$.
These do indeed match, subject to rounding, the first two moments of this data.

Exercise: The Inverse Gaussian Distribution with parameters μ and θ, has a mean of μ and a variance of μ^3/θ. Fit the Inverse Gaussian Distribution to this same data.
[Solution: Matching the mean and variance is equivalent to matching the first two moments.
$\mu = 62$.
$\mu^3/\theta = 9198.2 - 62^2 = 5354.2$
$\Rightarrow \theta = 62^3/ 5354.2 = 44.51$.]

Problems:

Use the following information for the next two questions.
Over one year, the following claim frequency observations were made for a group of 1,000 policies:

# of Claims	# of Policies
0	800
1	180
2	19
3	1

2.1 (2 points) You fit a Poisson Distribution via the method of moments.
Estimate the probability that a policy chosen at random will have more than 1 claim next year.
(A) Less than 1.6%
(B) At least 1.6%, but less than 1.8%
(C) At least 1.8%, but less than 2.0%
(D) At least 2.0%, but less than 2.2%
(E) At least 2.2%

2.2 (2 points) You fit a Binomial Distribution with m = 3 via the method of moments.
Estimate the probability that a policy chosen at random will have more than 1 claim next year.
(A) Less than 1.6%
(B) At least 1.6%, but less than 1.8%
(C) At least 1.8%, but less than 2.0%
(D) At least 2.0%, but less than 2.2%
(E) At least 2.2%

2.3 (1 point) The following five losses have been observed:

$50, $80, $100, $125, $220.

An exponential distribution $F(x) = 1 - e^{-x/\theta}$ is fit to this data by the method of moments.
What is the value of the fitted parameter θ?
A. less than 100
B. at least 100 but less than 110
C. at least 110 but less than 120
D. at least 120 but less than 130
E. at least 130

2.4 (3 points) The following five losses have been observed:

$500, $1,000, $1,500, $2,500, $4,500.

Use the method of moments to fit a LogNormal Distribution.
Use this LogNormal Distribution to estimate the probability that a loss will exceed $4,500.

A. Less than 5%
B. At least 5% but less than 6%
C. At least 6% but less than 7%
D. At least 7% but less than 8%
E. At least 8%

Use the following information for the next two questions.
Over one year, the following claim frequency observations were made for a group of 13,000 policies, where n_i is the number of claims observed for policy i:

$\sum n_i = 671$. $\sum n_i^2 = 822$.

2.5 (2 points) You fit a Poisson Distribution via the method of moments. Estimate the probability that a policy chosen at random will have at least one claim next year.

A. 4.4% B. 4.6% C. 4.8% D. 5.0% E. 5.2%

2.6 (3 points) You fit a Negative Binomial Distribution via the method of moments.
Estimate the probability that a policy chosen at random will have at least one claim next year.

A. 4.4% B. 4.6% C. 4.8% D. 5.0% E. 5.2%

The following information should be used to answer the next two questions:
10 Claims have been observed: 1500, 5500 3000, 3300, 2300, 6000, 5000, 4000, 3800, 2500. The underlying distribution is assumed to be Gamma, with parameters α and θ unknown.

2.7 (2 points) Determine the method of moments estimator of θ.

A. 480 B. 500 C. 520 D. 540 E. 560

2.8 (1 point) Determine the method of moments estimator of α

A. 3 B. 4 C. 5 D. 6 E. 7

Use the following information to answer the following four questions:
You observe the following five claims: 410, 1924, 2635, 4548, 6142.

2.9 (1 point) Using the method of moments, a LogNormal distribution is fit to this data.
What is the value of the fitted μ parameter?
A. less than 7.8
B. at least 7.8 but less than 7.9
C. at least 7.9 but less than 8.0
D. at least 8.0 but less than 8.1
E. at least 8.1

2.10 (2 points) Using the method of moments, a LogNormal distribution is fit to this data. What is the value of the fitted σ parameter?
A. less than .6
B. at least .6 but less than .7
C. at least .7 but less than .8
D. at least .8 but less than .9
E. at least .9

2.11 (1 point) Using the method of moments, a Normal distribution is fit to the natural logarithms of this data. What is the value of the fitted μ parameter?
A. less than 7.8
B. at least 7.8 but less than 7.9
C. at least 7.9 but less than 8.0
D. at least 8.0 but less than 8.1
E. at least 8.1

2.12 (2 points) Using the method of moments, a Normal distribution is fit to the natural logarithms of this data. What is the value of the fitted σ parameter?
A. less than .6
B. at least .6 but less than .7
C. at least .7 but less than .8
D. at least .8 but less than .9
E. at least .9

2.13 (1 point) You observe the following 6 claims: 16, 15, 10, 17, 2, 3.

A Distribution: $F(x) = 1 - e^{-\lambda x}$, $x > 0$, is fit to this data via the Method of Moments.

Determine the value of λ.

A. less than 0.06

B. at least 0.06 but less than 0.07

C. at least 0.07 but less than 0.08

D. at least 0.08 but less than 0.09

E. at least 0.09

Use the following information in the next six questions:

You observe the following 10 claims: 1729, 101, 384, 121, 880, 3043, 205, 132, 214, 82.

2.14 (2 points) You fit this data via the method of moments to a Pareto Distribution.

Determine α.

A. 4.0 B. 4.5 C. 5.0 D. 5.5 E. 6.0

2.15 (1 point) You fit this data via the method of moments to a Pareto Distribution.

Determine θ.

A. 2400 B. 2500 C. 2600 D. 2700 E. 2800

2.16 (1 point) The Inverse Gaussian Distribution with parameters μ and θ, has a mean of μ and a variance of μ^3/θ. You fit this data via the method of moments to an Inverse Gaussian Distribution.

In which of the following intervals is μ?

A. less than 400

B. at least 400 but less than 500

C. at least 500 but less than 600

D. at least 600 but less than 700

E. at least 700

2.17 (1 point) You fit this data via the method of moments to an Inverse Gaussian Distribution.

In which of the following intervals is θ?

A. less than 350

B. at least 350 but less than 360

C. at least 360 but less than 370

D. at least 370 but less than 380

E. at least 380

2.18 (1 point) The Inverse Gamma Distribution with parameters α and θ, has a mean of $\theta/(\alpha-1)$ and a second moment of $\frac{\theta^2}{(\alpha-1)(\alpha-2)}$. You fit this data via the method of moments to an Inverse Gamma Distribution. In which of the following intervals is α?

A. less than 1
B. at least 1 but less than 2
C. at least 2 but less than 3
D. at least 3 but less than 4
E. at least 4

2.19 (1 point) You fit this data via the method of moments to an Inverse Gamma Distribution.
In which of the following intervals is θ?

A. less than 1000
B. at least 1000 but less than 1100
C. at least 1100 but less than 1200
D. at least 1200 but less than 1300
E. at least 1300

2.20 (2 points) For a sample of 100 claims $x_1, x_2, ..., x_{100}$, you are given:

(i) $\sum x_i = 48{,}600$ and $\sum x_i^2 = 44{,}574{,}802$.

(ii) Claims are assumed to follow a lognormal distribution with parameters μ and σ.

(iii) μ and σ are estimated using the method of moments.

What is the fitted value of μ?

(A) Less than 5.00
(B) At least 5.00, but less than 5.25
(C) At least 5.25, but less than 5.50
(D) At least 5.50, but less than 5.75
(E) At least 5.75

2.21 (3 points) You observe the following 10 sizes of loss:

513 119 310 152 472 293 647 342 964 829

A distribution: $F(x) = 1 - (100/x)^q$, $x > 100$, is fit to this data via the Method of Moments.

Determine the value of q.

A. less than 1.1
B. at least 1.2 but less than 1.3
C. at least 1.3 but less than 1.4
D. at least 1.4 but less than 1.5
E. at least 1.5

Use the following information in the next two questions:
You observe the following 10 values: .21, .40, .14, .65, .53, .92, .30, .44, .76, .07.
The underlying distribution is a Beta Distribution as per Loss Models, with $\theta = 1$.
The mean of such a Beta Distribution is: a/(a+b).
The second moment of such a Beta Distribution is: $\frac{a\,(a+1)}{(a+b)\,(a+b+1)}$.

The parameters a and b are fit to this data via the Method of Moments.

2.22 (2 points) In which interval is the fitted a?
A. less than 1.0
B. at least 1.0 but less than 1.2
C. at least 1.2 but less than 1.4
D. at least 1.4 but less than 1.6
E. at least 1.6

2.23 (2 points) In which interval is the fitted b?
A. less than 1.0
B. at least 1.0 but less than 1.2
C. at least 1.2 but less than 1.4
D. at least 1.4 but less than 1.6
E. at least 1.6

2.24. (2 points) Loss sizes of more than 1000 follow a Single Parameter Pareto Distribution.
A sample of 10 loss sizes is as follows:
1200 1500 1500 1800 2000 2000 3000 5000 10,000 25,000
Calculate the method of moments estimate for the parameter α.
A. Less than 1.1
B. At least 1.1, but less than 1.2
C. At least 1.2, but less than 1.3
D. At least 1.3, but less than 1.4
E. At least 1.4

2.25 (2 points) The number of claims by policy for 10 polices were:

Policy:	A	B	C	D	E	F	G	H	I	J
Claims:	0	2	1	0	1	1	0	3	0	0

A Negative Binomial distribution is fit to this data via the method of moments.
What is the value of the fitted parameter r of the Negative Binomial distribution?
A. 1 B. 2 C. 3 D. 4 E. 5

2.26 (4, 5/86, Q.55) (2 points) $X_1, X_2,...X_n$ is an independent sample drawn from a lognormal distribution with parameters μ and σ^2.

Let $\overline{X} = (1/n) \sum_{i=1}^{n} X_i$ $\qquad S^2 = (1/(n-1)) \sum_{i=1}^{n} (X_i - \overline{X})^2$

In terms of $\overline{X}$ and S^2 obtain estimators for μ and σ^2 using the method of moments.

A. $\mu = \overline{X}$; $\sigma^2 = S^2 / \overline{X}$

B. $\mu = \ln \overline{X}$; $\sigma^2 = \ln S^2$

C. $\mu = \overline{X}$; $\sigma^2 = \ln (S^2 + \overline{X}^2)$

D. $\mu = .5 \ln (\overline{X}^3 / (\overline{X} + S^2))$; $\sigma^2 = \ln (S^2 - \overline{X}^2)$

E. $\mu = \ln \overline{X} - .5 \ln (1 + S^2/\overline{X}^2)$; $\sigma^2 = \ln (1 + S^2/\overline{X}^2)$

2.27 (4, 5/87, Q.60) (1 point) Using the method of moments what is an estimate of the mean of a lognormal distribution given the sample: 3, 4.5, 6, 6.25, 6.5, 6.75, 7, 7.5, 8.5, 10.

A. Less than 6
B. At least 6, but less than 6.25
C. At least 6.25, but less than 6.50
D. At least 6.50, but less than 6.75
E. 6.75 or more.

2.28 (4, 5/88, Q.53) (2 points) Given the distribution $f(x) = ax^{a-1}$, $0 < x < 1$, $0 < a < \infty$, and the sample 0.7, 0.14, 0.8, .09, 0.65; what is the method of moments estimate for a?

A. Less than 1.0
B. At least 1.0, but less than 1.3
C. At least 1.3, but less than 1.6
D. At least 1.6, but less than 1.9
E. 1.9 or more

2.29 (4, 5/89, Q.45) (2 points) The number of claims X that a randomly selected policyholder has in a year follows a negative binomial distribution, as per Loss Models.
The following data was compiled for 10,000 policyholders in one year:

Number of Claims	Number of Policyholders
0	8,200
1	1,000
2	600
3	200

What is the method of moments estimate for the parameter β?

A. $\beta < .35$

B. $.35 \le \beta < .40$

C. $.40 \le \beta < .45$

D. $.45 \le \beta < .50$

E. $.50 \le \beta$

2.30 (2, 5/90, Q. 31) (1.7 points) Let X be a continuous random variable with density function $f(x ; \theta) = x^{(1-\theta)/\theta}/\theta$ for $0 < x < 1$, where $\theta > 0$.
What is the method-of-moments estimator of θ?

A. $(1 - \bar{X})/\bar{X}$

B. $(\bar{X} - 1)/\bar{X}$

C. $\bar{X}/(1- \bar{X})$

D. $\bar{X}/(\bar{X} - 1)$

E. $1/(1 + \bar{X})$

2.31 (4, 5/90, Q.24) (1 point) Assume that the number of claims for an insured has a Poisson distribution: $p(n) = e^{-\theta} \theta^n / n!$.
Using the observations 3, 1, 2, 1, taken from a random sample, what is the method of moments estimate, $\tilde{\theta}$, of θ?

A. $\tilde{\theta} < 1.40$

B. $1.40 \le \tilde{\theta} < 1.50$

C. $1.50 \le \tilde{\theta} < 1.60$

D. $1.60 \le \tilde{\theta} < 1.70$

E. $1.70 \le \tilde{\theta}$

2.32 (4, 5/90, Q.34) (2 points) The following observations: 1000, 850, 750, 1100, 1250, 900, are a random sample taken from a Gamma distribution with unknown parameters α and θ.
In what range does the method of moments estimator of α fall?
A. Less than 30
B. At least 30, but less than 40
C. At least 40, but less than 50
D. At least 50, but less than 60
E. 60 or more

2.33 (4, 5/91, Q.46) (2 points) The following sample of 10 claims is observed:
1500 6000 3500 3800 1800 5500 4800 4200 3900 3000.
The underlying distribution is assumed to be Gamma, with parameters α and θ unknown.
In what range does the method of moments estimators of θ fall?
A. Less than 250
B. At least 250, but less than 300
C. At least 300, but less than 350
D. At least 350, but less than 400
E. 400 or more

2.34 (4B, 5/92, Q.5) (2 points) You are given the following information:
- Number of large claims follows a Poisson distribution.
- Exposures are constant and there are no inflationary effects.
- In the past 5 years, the following number of large claims has occurred: 12, 15, 19, 11, 18

Estimate the probability that more than 25 large claims occur in one year.
(The Poisson distribution should be approximated by the Normal distribution.)
A. Less than .002
B. At least .002 but less than .003
C. At least .003 but less than .004
D. At least .004 but less than .005
E. At least .005

2.35 (4B, 5/92, Q.10) (3 points) You are given the following information:

- Losses follow a LogNormal distribution with parameters μ and σ.
- The following five losses have been observed: \$500, \$1,000, \$1,500, \$2,500, \$4,500.

Use the method of moments to fit a Normal Distribution to the natural logs of the loss sizes. Use the corresponding LogNormal Distribution to estimate the probability that a loss will exceed \$4,500.

A. Less than 0.03
B. At least 0.03 but less than 0.06
C. At least 0.06 but less than 0.09
D. At least 0.09 but less than 0.12
E. At least 0.12

2.36 (4B, 5/92, Q.26) (1 point) The random variable X has the density function with parameter β given by $f(x;\beta) = (1/\beta^2)$ x $\exp(-.5\ (x/\beta)^2)$; $x > 0$, $\beta > 0$.

Where $E[X] = (\beta/2)\sqrt{2\pi}$ and the variance of X is $2\beta^2 - (\pi/2)\beta^2$.

You are given the following observations of X: 4.9, 1.8, 3.4, 6.9, 4.0.
Determine the method of moments estimate of β.

A. Less than 3.00
B. At least 3.00 but less than 3.15
C. At least 3.15 but less than 3.30
D. At least 3.30 but less than 3.45
E. At least 3.45

2.37 (4B, 11/92, Q.13) (2 points) You are given the following information:

- The occurrence of hurricanes in a given year has a Poisson distribution.
- For the last 10 years, the following number of hurricanes has occurred:
  2, 4, 3, 8, 2, 7, 6, 3, 5, 2

Using the normal approximation to the Poisson, determine the probability of more than 10 hurricanes occurring in a single year.

A. Less than 0.0005
B. At least 0.0005 but less than 0.0025
C. At least 0.0025 but less than 0.0045
D. At least 0.0045 but less than 0.0065
E. At least 0.0065

2.38 (4B, 5/95, Q.5) (2 points) You are given the following:
• The random variable X has the density function

$f(x) = \alpha x^{\alpha-1}$, $0 < x < 1$, $\alpha > 0$.

• A random sample of three observations of X yields the values: 0.40, 0.70, 0.90.
Determine the method of moments estimate of α.
A. Less than 0.5
B. At least 0.5, but less than 1.5
C. At least 1.5, but less than 2.5
D. At least 2.5, but less than 3.5
E. At least 3.5

<u>2.39</u> **(2, 2/96, Q.26)** (1.7 points) Let $X_1,..., X_n$ be a random sample from a discrete distribution with probability function: $p(1) = \theta$, $p(2) = \theta$, and $p(3) = 1 - 2\theta$, where $0 < \theta < 1/2$.
Determine the method of moments estimator of θ .
A. $(3 - \bar{X})/3$ B. $(\bar{X} - 1)/4$ C. $(2\bar{X} - 3)/6$ D. $\bar{X}$ E. $\bar{X}/2$

2.40 (4B, 5/96, Q.4) (2 points) You are given the following:
• The random variable X has the density function

$f(x) = 2(\theta-x)/\theta^2$, $0 < x < \theta$.

• A random sample of two observations of X yields the values 0.50 and 0.90.
Determine the method of moments estimator of θ.
A. Less than 0.45
B. At least 0.45, but less than 0.95
C. At least 0.95, but less than 1.45
D. At least 1.45, but less than 1.95
E. At least 1.95

2.41 (4B, 11/97, Q.18) (2 points) You are given the following:
• The random variable X has the density function $f(x) = \alpha\, x^{-(\alpha+1)}$, $1 < x < \infty$, $\alpha > 1$.
• A random sample is taken of the random variable X.
Determine the limit of the method of moments estimator of α, as the sample mean goes to infinity.
A. 0 B. 1/2 C. 1 D. 2 E. ∞

2.42 (4B, 5/98, Q.5) (2 points) You are given the following:

- The random variables X has the density function
  $f(x) = \alpha\ (x+1)^{-(\alpha+1)}$, $0 < x < \infty$, $\alpha > 0$.
- A random sample of size n is taken of the random variable X.
- The values in the random sample totals $n\mu$.

Assuming $\alpha > 1$, determine the method of moments estimator of α.

A. μ B. $\mu / (\mu - 1)$ C. $\mu / (\mu + 1)$ D. $(\mu - 1) / \mu$ E. $(\mu + 1) / \mu$

2.43 (IOA 101, 4/00, Q.9) (3.75 points)
The discrete random variable X has the following probability function:
$P(X = i) = 0.2 + \alpha(i - 2)$: i = 0, 1, 2, 3, 4.
(i) (0.75 points) State the possible values that α can take.
(ii) (3 points) Given a random sample x_1 , x_2 , ..., x_n from this distribution, determine the method of moments estimate of α and show that this can result in inadmissible estimates (i.e. estimates outside the range of possible values of α).

2.44 (CAS3, 5/05, Q.19) (2.5 points) Four losses are observed from a Gamma distribution. The observed losses are: 200, 300, 350, and 450. Find the method of moments estimate for α.

A. 0.3 B. 1.2 C. 2.3 D. 6.7 E. 13.0

2.45 (CAS3, 5/06, Q.1) (2.5 points) The number of goals scored in a soccer game follows a Negative Binomial distribution. A random sample of 20 games produced the following distribution of the number of goals scored:

Goals Scored	Frequency
0	1
1	3
2	4
3	5
4	3
5	2
6	1
7	0
8	1

Use the sample data and the method of moments to estimate the parameter β of the Negative Binomial distribution.

A. Less than 0.25
B. At least 0.25, but less than 0.50
C. At least 0.50, but less than 0.75
D. At least 0.75, but less than 1.00
E. At least 1.00

2.46 (CAS3, 11/06, Q.3) (2.5 points) Claim sizes of 10 or greater are described by a single parameter Pareto distribution, with parameter α. A sample of claim sizes is as follows:

10 12 14 18 21 25

Calculate the method of moments estimate for α for this sample.

A. Less than 2.0
B. At least 2.0, but less than 2.1
C. At least 2.1, but less than 2.2
D. At least 2.2, but less than 2.3
E. At least 2.3

2.47 (CAS3, 11/07, Q.5) (2.5 points)
X is a two-parameter Pareto random variable with parameters θ and α.
A random sample from this distribution produces the following four claims:

- x_1 = 2,000
- x_2 = 17,000
- x_3 = 271,000
- x_4 = 10,000

Find the Method of Moments estimate for α.

A. Less than 2
B. At least 2, but less than 3
C. At least 3, but less than 4
D. At least 4, but less than 5
E. At least 5

2.48 (CAS3L, 5/09, Q.17) (2.5 points) A random variable, X, follows a lognormal distribution.
You are given a sample of size n and the following information:

$\Sigma\, x_i / n = 1.8682$

$\Sigma\, x_i^2 / n = 4.4817$

Use the method of moments to estimate the lognormal parameter σ.

A. Less than 0.4
B. At least 0.4, but less than 0.8
C. At least 0.8, but less than 1.2
D. At least 1.2 but less than 1.6
E. At least 1.6

2.49 (CAS3L, 11/09, Q.17) (2.5 points) You are creating a model to describe exam progress.
You are given the following information:

- Let X be the number of exams passed in a given year.
- The probability mass function is defined as follows:

  P(X = 0) = 1 - p - q
  P(X = 1) = p
  P(X = 2) = q
- Over the last 5 years, you observe the following values of X:

  0 0 1 2 2

Calculate the method of moments estimate of p.

A. Less than 0.15
B. At least 0.15, but less than 0.21
C. At least 0.21, but less than 0.27
D. At least 0.27, but less than 0.33
E. At least 0.33

Section 3, Variance of Estimated Parameters, Method of Moments

One is interested in how precise is the estimate of a parameter.
One important measure of this, is the variance of the estimate.

Variance of an Estimated Mean:

Let us assume we have X_1, X_2 and X_3, three independent, identically distributed variables.
Since they are independent, their variances add, and since they are identical they each have the same variance, Var[X]:
$\text{Var}[X_1 + X_2 + X_3] = \text{Var}[X_1] + \text{Var}[X_2] + \text{Var}[X_3] = 3\text{Var}[X]$.

Let the estimated mean be: $\overline{X} = (X_1 + X_2 + X_3)/3$.

$\text{Var}[\overline{X}] = \text{Var}[(X_1 + X_2 + X_3)/3] = \text{Var}[X_1 + X_2 + X_3]/3^2 = (3\text{Var}[X])/3^2 = \text{Var}[X]/3$.

Exercise: We generate four independent random variables, each from a Poisson with $\lambda = 5.6$.
What is the variance of their average?
[Solution: $\text{Var}[\overline{X}] = \text{Var}[(X_1 + X_2 + X_3 + X_4)/4] = \text{Var}[X_1 + X_2 + X_3 + X_4]/4^2 = (4\text{Var}[X])/4^2 = \text{Var}[X]/4 = 5.6/4 = 1.4$.]

For n independent, identically distributed variables:

$$\text{Var}[\overline{X}] = \text{Var}[(1/n)\sum_{i=1}^{n} X_i] = \text{Var}[\sum_{i=1}^{n} X_i]/n^2 = \sum_{i=1}^{n} \text{Var}[X_i]/n^2 = n\,\text{Var}[X]/n^2 = \text{Var}[X]/n.$$

$$\mathbf{Var}[\overline{X}] = \mathbf{Var}[X] / n.$$

The variance of an average declines as 1/(the number of data points).[7]

Variance of Estimated Parameters:

Previously, we fit a Poisson Distribution to 23,872 observations via the Method of Moments and obtained a point estimate $\lambda = .3346$. Then one might ask how reliable this estimate is, assuming this data actually came from a Poisson Distribution. In other words, what is the variance one would expect in this estimate solely due to random fluctuation in the data set?

---

[7] *The skewness of $\overline{X}$ goes down as $1/\sqrt{n}$. Excess = kurtosis - 3. The excess of $\overline{X}$ goes down as 1/n.*
See equation 5.14.2 in Statistical Methods by Snedecor and Cochran.

Using the Method of Moments, we set $\hat{\lambda} = \bar{X}$. Therefore $\text{Var}[\hat{\lambda}] = \text{Var}[\bar{X}]$.

Thus in this case we have reduced the problem of calculating the variance of the estimated λ to that of estimating the variance of the estimated mean.
For X from a Poisson Distribution, with $\lambda = .3346$: $\text{Var}(X) = \lambda = .3346$.

Therefore, $\text{Var}[\hat{\lambda}] = \text{Var}[\bar{X}] = \text{Var}[X]/n = \hat{\lambda} / n = .3346 / 23872 = .00001402$.

The standard deviation of the estimated λ is thus .00374. Using plus or minus 1.96 standard deviations, an approximate 95% confidence interval for $\hat{\lambda}$ is $.3346 \pm .0073$.[8]

The larger the data set, the larger n, and the smaller the variance of the estimate of λ.

Exercise: A Binomial Distribution with fixed m = 6 has been fit to 1000 observations via the Method of Moments. The resulting estimate of $\hat{q} = .238$. What is an approximate 95% confidence interval for $\hat{q}$?

[Solution: $\hat{q} = \bar{X}/6$. $\text{Var}[\hat{q}] = \text{Var}[\bar{X}]/36 = (\text{Var}[X]/1000)/36 = \text{Var}[X]/36000 =$
$6(q)(1-q) / 36000 = (0.238)(1 - 0.238)/6000 = .00003023$. The Standard Deviation is:
$\sqrt{0.00003023} = .0055$. Thus using plus or minus 1.96 standard deviations, an approximate 95% confidence interval for $\hat{q}$ is: 0.238 ± 0.011.]

Exercise: Assume we have fit 23,872 observations via the Method of Moments to a Negative Binomial Distribution with fixed r = 1.5 and obtained a point estimate $\hat{\beta} = .223$.

What is an approximate 95% confidence interval for $\hat{\beta}$?

[Solution: $\hat{\beta} = \bar{X}/1.5$. $\text{Var}[\hat{\beta}] = \text{Var}[\bar{X}]/2.25 = \dfrac{\text{Var}[X]/ 23{,}872}{2.25} = \dfrac{1.5\ \beta\ (1+\beta)}{(2.25)\ (23{,}872)} =$

$\dfrac{(0.223)\ (1.223)}{(1.5)\ (23{,}872)} = 0.00000762$. Standard Deviation is: $\sqrt{0.00000762} = .00276$.

Thus an approximate 95% confidence interval for $\hat{\beta}$ is: 0.223 ± 0.005.

<u>Comment</u>: The confidence interval is so narrow due to the large amount of data. Its width goes down as the inverse of the square root of the amount of data.]

---

[8] Since for a Poisson, the Method of Moments is equal to the Method of Maximum Likelihood, this is also the interval estimate from the Method of Maximum Likelihood.

Exercise: A data set of 10 losses is: 2, 7, 11, 27, 29, 47, 52, 75, 110, 260.

$\bar{X}$ = 62. $E[X^2]$ = 9198.2.

Fit a Pareto Distribution with α = 4 (fixed) and θ unknown via method of moments to this data.

Determine an approximate 90% confidence interval for θ.

[Solution: The mean of this Pareto is $\theta/(\alpha-1) = \theta/3$. Since there is one parameter, the method of moments involves matching the mean: $\bar{X} = \theta/3$.

$\hat{\theta} = 3\bar{X} = (3)(62) = 186.$

$\hat{\theta} = 3\bar{X}. \Rightarrow \text{Var}(\hat{\theta}) = 3^2\ \text{Var}(\bar{X}) = 9\ \text{Var}(X) / n.$

For X from a Pareto Distribution, with α = 2.5: Var(X) = 2nd moment - mean^2 =

$$\frac{2\,\theta^2}{(\alpha-1)(\alpha-2)} - \left(\frac{\alpha\,\theta}{\alpha-1}\right)^2 = \frac{\alpha\,\theta^2}{(\alpha-1)^2(\alpha-2)} = \theta^2\frac{4}{(3^2)\,(2)} = (2/9)\theta^2.$$

Thus, $\text{Var}(\hat{\theta}) = 9\ \text{Var}(X) / n = (9)(2/9)\theta^2/10 = 0.2\theta^2 = (0.2)(186^2) = 6919.$

Thus an approximate 90% confidence interval for θ is:

$\hat{\theta} = 186 \pm (1.645)(\sqrt{6919}) = 186 \pm 137.$

<u>Comment</u>: The confidence interval is very wide because we have only 10 losses in the data set.]

Problems:

3.1 (3 points) You are given the following:

- The random variable X has the density function: $f(x) = (1/\theta)\, e^{-x/\theta}$, x >0.
- A random sample of 6 observations of X yields the values: 10, 12, 15, 17, 19, 23.

Estimate the variance of the method of moments estimator of θ.

A. Less than 30
B. At least 30, but less than 35
C. At least 35, but less than 40
D. At least 40, but less than 45
E. At least 45

Use the following information for the next three questions:
One has observed the following distribution of insureds by number of claims:

Number of Claims	0	1	2	3	4	5&+	All
Number of Insureds	490	711	572	212	15	0	2000

3.2 (2 points) A Binomial Distribution with m = 4 is fit via the Method of Moments.
Which of the following is an approximate 95% confidence interval for q?

A. [.317, .321] B. [.315, .323] C. [.313, .325]
D. [.311, .327] E. [.309, .329]

3.3 (2 points) A Poisson Distribution is fit via the Method of Moments.
Which of the following is an approximate 95% confidence interval for λ?

A. [1.251, 1.301] B. [1.226, 1.326] C. [1.201, 1.351]
D. [1.176, 1.376] E. [1.151, 1.401]

3.4 (2 points) A Negative Binomial Distribution with r = 2 is fit via the Method of Moments.
Which of the following is an approximate 95% confidence interval for β?

A. [.636, .640] B. [.626, .650] C. [.606, .670]
D. [.576, .700] E. [.526, .750]

Use the following information for the next 2 questions:

You observe 1000 individual claims, which you assume follow a Pareto Distribution.

You are given that:

$$\sum_{i=1}^{1000} x_i = 72{,}134.$$

3.5 (1 point) Assuming $\alpha = 2.5$, estimate θ using the Method of Moments.

A. Less than 100

B. At least 100, but less than 105

C. At least 105, but less than 110

D. At least 110, but less than 115

E. At least 115

3.6 (2 points) Estimate the variance of the estimate of θ in the previous question.

A. Less than 50

B. At least 50, but less than 52

C. At least 52, but less than 54

D. At least 54, but less than 56

E. At least 56

Use the following information for the next 2 questions:
• You observe 100,000 losses for a total of $873 million.
• Based on having analyzed similar data from previous years, you believe these losses were drawn from a Weibull Distribution, as per <u>Loss Models</u>, with parameters τ = 1/4 and θ unknown.

3.7 (1 point) Estimate θ using the method of moments.
A. Less than 370
B. At least 370, but less than 380
C. At least 380, but less than 390
D. At least 390, but less than 400
E. At least 400

3.8 (2 points) What is the standard deviation of the estimate of θ?
A. Less than 6
B. At least 6, but less than 7
C. At least 7, but less than 8
D. At least 8, but less than 9
E. At least 9

Use the following information for the next 6 questions:
During a year, 10,000 insureds have a total of 4200 claims.

3.9 (1 point) You fit a Binomial Distribution with m = 10 via the method of moments. What is the fitted q?
(A) Less than .04
(B) At least .04, but less than .05
(C) At least .05, but less than .06
(D) At least .06, but less than .07
(E) At least .07

3.10 (2 points) What is the standard deviation of $\hat{q}$?
(A) Less than .0004
(B) At least .0004, but less than .0005
(C) At least .0005, but less than .0006
(D) At least .0006, but less than .0007
(E) At least .0007

3.11 (1 point) You fit a Poisson Distribution via the method of moments. What is the fitted λ?
(A) Less than .4
(B) At least .4, but less than .5
(C) At least .5, but less than .6
(D) At least .6, but less than .7
(E) At least .7

3.12 (1 point) What is the standard deviation of $\hat{\lambda}$?
(A) Less than .007
(B) At least .007, but less than .008
(C) At least .008, but less than .009
(D) At least .009, but less than .010
(E) At least .010

3.13 (1 point) You fit a Negative Binomial Distribution with r = 3 via the method of moments. What is the fitted β?
(A) .14 (B) .16 (C) .18 (D) .20 (E) .22

3.14 (2 points) What is the standard deviation of $\hat{\beta}$?
(A) Less than .001
(B) At least .001, but less than .002
(C) At least .002, but less than .003
(D) At least .003, but less than .004
(E) At least .004

Use the following information for the next 4 questions:
You observe 10 individual claims: 1534, 567, 383, 434, 563, 132, 1262, 748, 299, 516
which you assume follow a Gamma Distribution.

3.15 (1 point) Assuming $\alpha = 3$, estimate θ using the Method of Moments.
A. Less than 190
B. At least 190, but less than 200
C. At least 200, but less than 210
D. At least 210, but less than 220
E. At least 220

3.16 (3 points) Estimate the standard deviation of the estimate of θ in the previous question.
A. Less than 20
B. At least 20, but less than 25
C. At least 25, but less than 30
D. At least 30, but less than 35
E. At least 35

3.17 (1 point) Assuming $\theta = 200$, estimate α using the Method of Moments.
A. Less than 3.0
B. At least 3.0, but less than 3.2
C. At least 3.2, but less than 3.4
D. At least 3.4, but less than 3.6
E. At least 3.6

3.18 (2 points) Estimate the variance of the estimate of α in the previous question.
A. Less than .4
B. At least .4, but less than .5
C. At least .5, but less than .6
D. At least .6, but less than .7
E. At least .7

Use the following information to answer the next two questions:

10,000 individual claims are observed with a total of $1,050,000 in losses.

3.19 (1 point) An exponential distribution $F(x) = 1 - e^{-x/\theta}$ is fit to this data via the method of moments. What is the resulting estimate of θ?

A. less than 110
B. at least 110 but less than 115
C. at least 115 but less than 120
D. at least 120 but less than 125
E. at least 125

3.20 (1 point) What is the standard deviation of the estimate of θ in the prior question?

A. less than 0.9
B. at least 0.9 but less than 1.0
C. at least 1.0 but less than 1.1
D. at least 1.1 but less than 1.2
E. at least 1.2

3.21 (2, 5/85, Q. 31) (1.5 points) Some scientists believe that Drug X would benefit about half of all people with a certain blood disorder. To estimate the proportion, p, of patients who would benefit from taking Drug X, the scientists will administer it to a random sample of patients who have the blood disorder. The estimate of p will be $\hat{p}$, the proportion of patients in the sample who benefit from having taken the drug. Which of the following is closest to the minimum sample size that guarantees $P[|p - \hat{p}| \le .03] \ge .95$?

A. 748 B. 1,068 C. 1,503 D. 2,056 E. 2,401

3.22 (4B, 11/95,Q.17) (3 points) You are given the following:

- Losses follow a Pareto distribution, with parameters θ (unknown) and $\alpha = 3$.
- 300 losses have been observed.

Determine the variance of the method of moments estimator of θ.

A. $.0025\ \theta^2$ B. $.0033\ \theta^2$ C. $.0050\ \theta^2$ D. $.0100\ \theta^2$ E. $.0133\ \theta^2$

3.23 (IOA 101, 9/00, Q.5) (1.5 points) The number of claims which arise under a policy of a particular type in a year is to be modeled as a Poisson (λ) random variable.
A random sample of 500 such policies gave rise to a total of 84 claims in 1999.
Calculate a 95% confidence interval for λ.