

Average Breakdown on Exam STAM:

Frequency Distributions	5%
Loss Distributions	8%
Aggregate Distributions	6%
Risk Measures	3%
Fitting Frequency Distributions	4%
Fitting Loss Distributions	20%
Credibility	27%
P&C Ratemaking and Reserving	24%
Individual Health Insurance	3%

The number of questions by topic varies from exam to exam.

Exam **STAM**

Howard Mahler

Frequency Distributions

The slides are in the same order as the sections of my study guide.

Section #	Section Name
1	Introduction
2	Basic Concepts
3	Binomial Distribution
4	Poisson Distribution
5	Geometric Distribution
6	Negative Binomial Distribution
7	Normal Approximation
8	<i>Skewness</i>
9	Probability Generating Functions
10	<i>Factorial Moments</i>
11	(a, b, 0) Class of Distributions
12	<i>Accident Profiles</i>
13	Zero-Truncated Distributions
14	Zero-Modified Distributions
15	<i>Compound Frequency Distributions</i>
16	Moments of Compound Distributions
17	Mixed Frequency Distributions
18	Gamma Function
19	Gamma-Poisson Frequency Process
20	<i>Tails of Frequency Distributions</i>

At the end, there are some additional questions for study.

After the slides are my section of important ideas and formulas.

Introduction

Sections 2 to 7 cover preliminary material which you should know, even though it is not on the syllabus of STAM.

While this material should not be tested directly, it will help you to answer exam questions.

I would review this material to the extent you need to.

Basic Concepts

A Frequency Distribution is discrete.

It has probability on the nonnegative integers.

The probabilities add to one.

Number of Claims	Probability Density Function	Probability times # of Claims	Probability times Square of # of Claims
0	0.1	0	0
1	0.2	0.2	0.2
2	0	0	0
3	0.1	0.3	0.9
4	0	0	0
5	0	0	0
6	0.1	0.6	3.6
7	0	0	0
8	0	0	0
9	0.1	0.9	8.1
10	0.3	3	30
11	0.1	1.1	12.1
Sum	1	6.1	54.9

Average of X = 1st moment about the origin =
6.1 = Mean.

Average of X^2 = 2nd moment about the origin = 54.9.

Variance = $54.9 - 6.1^2 = 17.7$.

Binomial Distribution:

Support: $x = 0, 1, 2, 3, \dots, m$. Finite support.

Parameters: $1 > q > 0$, $m \geq 1$, integer.

$$f(x) = \frac{m!}{x! (m-x)!} q^x (1-q)^{m-x}, 0 \leq x \leq m.$$

Mean = mq

Variance = $mq(1-q)$

For example, for $m = 5$ and $q = 0.1$:

$$f(2) = \frac{5!}{2! 3!} 0.1^2 0.9^{5-2} = (10) 0.1^2 0.9^3 = 0.0729.$$

$f(x) \Leftrightarrow p_k$.

Binomial Distribution for $m = 1$ is a Bernoulli Distribution.

Bernoulli Distribution: $f(1) = q$, $f(0) = 1 - q$.

The sum of m independent, identically distributed Bernoulli Distributions is a Binomial Distribution.

The sum of two independent Binomials with the same q is another Binomial:

$$\text{Binom}(m_1, q) + \text{Binom}(m_2, q) = \text{Binom}(m_1 + m_2, q).$$

TI-30XS **Multiview** calculator saves time doing repeated calculations using the same formula.

For example constructing a table of the densities of a Binomial distribution, with $m = 5$ and $q = 0.1$:

$$f(x) = \binom{5}{x} 0.1^x 0.9^{5-x}.$$

table

$$y = (5 \text{ nCr } x) * .1^x * .9^{(5-x)}$$

Enter

Start = 0

Step = 1

Auto

OK

$$x = 0 \quad y = 0.59049$$

$$x = 1 \quad y = 0.32805$$

$$x = 2 \quad y = 0.07290$$

$$x = 3 \quad y = 0.00810$$

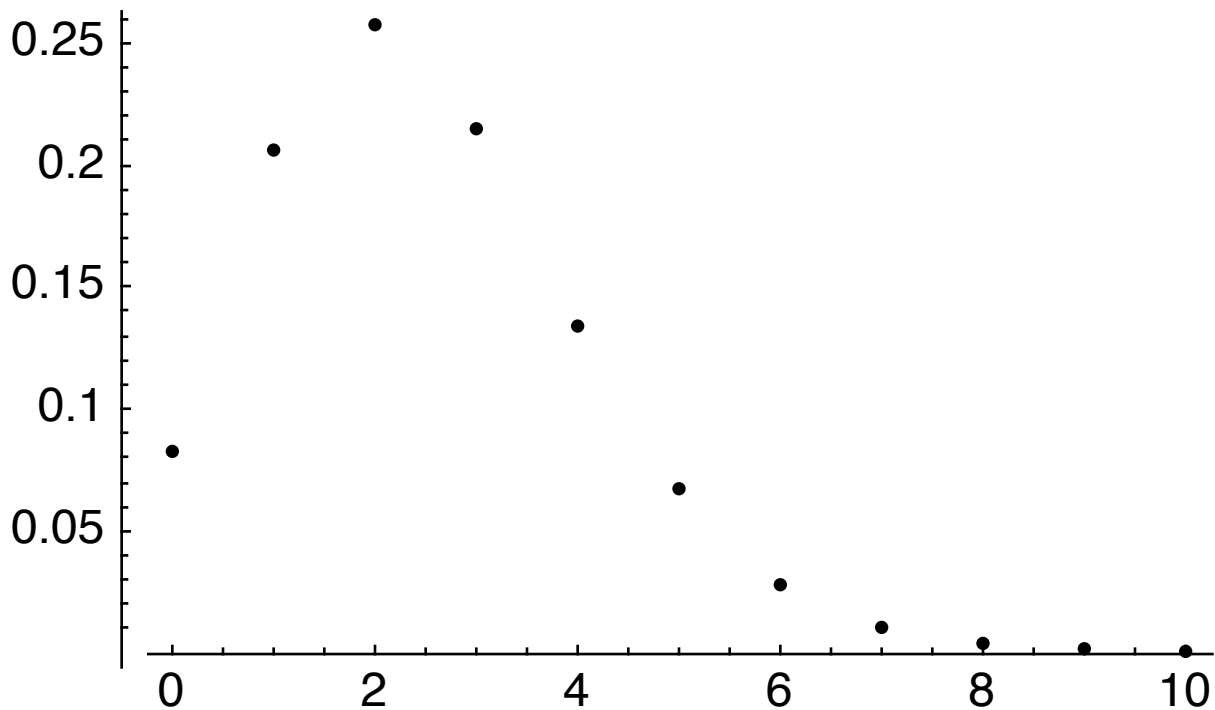
$$x = 4 \quad y = 0.00045$$

$$x = 5 \quad y = 0.00001$$

Poisson Distribution:

$$f(x) = \frac{\lambda^x e^{-\lambda}}{x!}, \quad x \geq 0 \text{ (infinite support)}$$

$$\text{If } \lambda = 2.5, f(4) = \frac{\lambda^4 e^{-\lambda}}{4!} = \frac{2.5^4 e^{-2.5}}{4!} = 13.36\%.$$



Mean = λ Variance = λ Mean = Variance.

A Poisson Distribution \Leftrightarrow

a constant independent claim intensity.

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

Frequency Poisson and severity independent of frequency \Rightarrow

the number of claims above a certain amount is also a Poisson.

This is called “thinning**” a Poisson.**

The small claims are also Poisson.

The number of small and large claims are **independent** of each other.

4.6. The male drivers in the State of Grace each have their annual claim frequency given by a Poisson distribution with parameter equal to 0.05. The female drivers in the State of Grace each have their annual claim frequency given by a Poisson distribution with parameter equal to 0.03. You insure in the State of Grace 20 male drivers and 10 female drivers. Assume the claim frequency distributions of the individual drivers are independent. What is the chance of observing 3 claims in a year?

4.6. E. The sum of the claims from 20 males is Poisson with mean: $(20)(0.05) = 1.0$.

The sum of the claims from 10 females is Poisson with mean: $(10)(0.03) = 0.3$.

The sum of the claims from everyone is Poisson with mean: $1.0 + 0.3 = 1.3$.

The chance of 3 claims is: $(1.3^3) e^{-1.3} / 3! = \mathbf{9.98\%}$.

Next Q. 4.12.

4.12. Frequency follows a Poisson Distribution with $\lambda = 7$.

20% of losses are of size greater than \$50,000.

Frequency and severity are independent.

Let N be the number of losses of size greater than \$50,000. What is the probability that $N = 3$?

4.12. D. Number of large losses follows a Poisson Distribution with $\lambda = (20\%)(7) = 1.4$.

$$f(3) = \frac{1.4^3 e^{-1.4}}{3!} = \mathbf{11.3\%}.$$

Comment: Number of small losses follows a Poisson Distribution with $\lambda = (80\%)(7) = 5.6$.

Number of small losses is **independent** of the number of large losses!

Geometric Distribution

$$f(x) = \frac{\beta^x}{(1+\beta)^{x+1}}, x = 0, 1, 2, 3, \dots$$

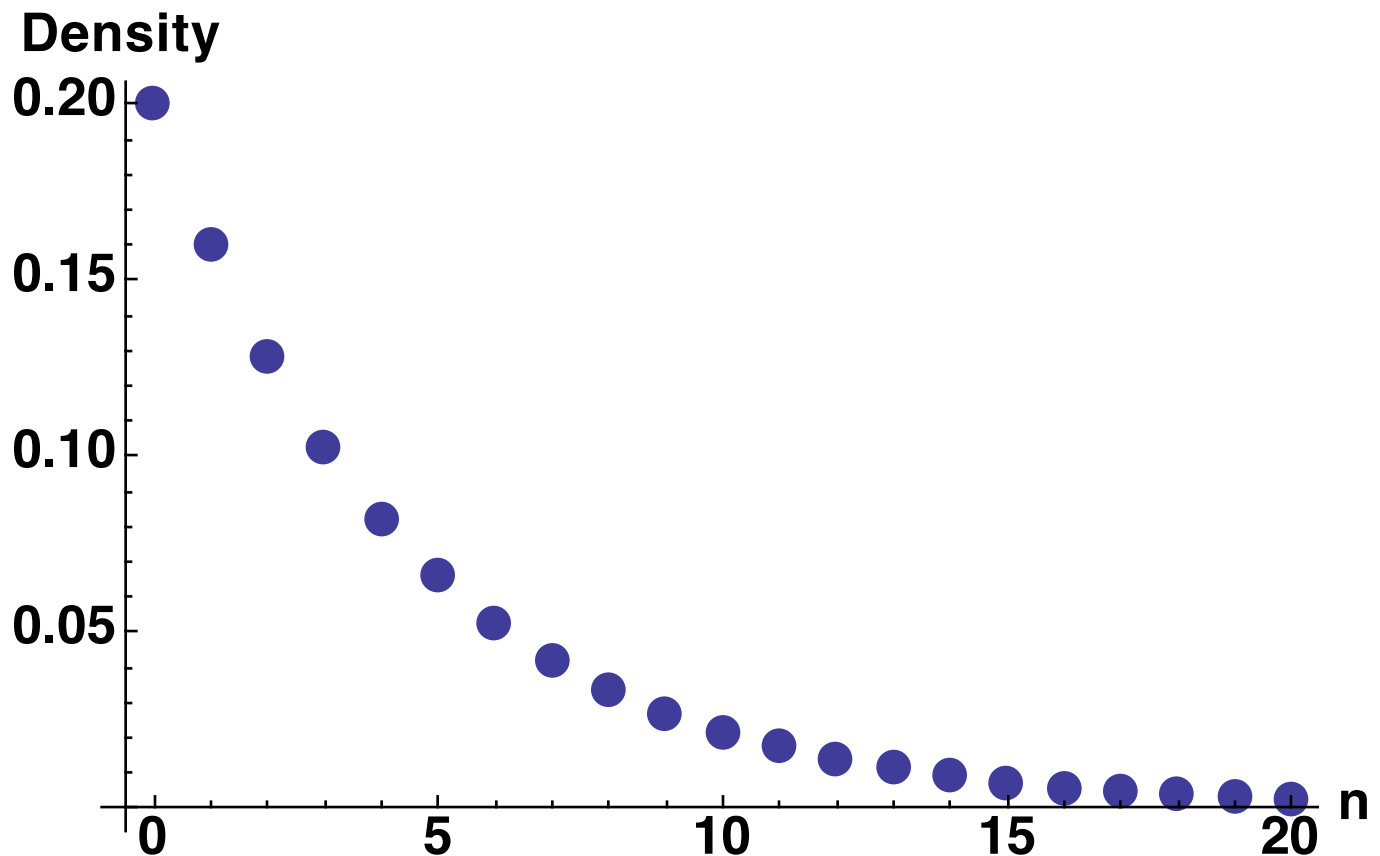
Mean = β . Variance = $\beta(1+\beta)$. Mode = 0.

$$f(0) = \frac{1}{1 + \beta}.$$

$$f(n+1) = f(n) \frac{\beta}{1 + \beta} < f(n).$$

$$f(n+1) = f(n) \frac{\beta}{1 + \beta} < f(n).$$

A Geometric Distribution with $\beta = 4$:



Geometric Distribution \Leftrightarrow Exponential Distribution.

5.33. 1, 5/00, Q.36. In modeling the number of claims filed by an individual under an automobile policy during a three-year period, an actuary makes the simplifying assumption that for all integers $n \geq 0$, $p_{n+1} = p_n/5$, where p_n represents the probability that the policyholder files n claims during the period. Under this assumption, what is the probability that a policyholder files more than one claim during the period?

1, 5/00, Q.36. A.

The densities are declining geometrically.

Therefore, this is a Geometric Distribution, with

$$\frac{\beta}{1 + \beta} = 1/5. \Rightarrow \beta = 1/4.$$

Prob[more than one claim] = $1 - f(0) - f(1) =$

$$1 - \frac{1}{1 + \beta} - \frac{\beta}{(1 + \beta)^2} = 1 - 4/5 - 4/25 = \mathbf{4\%}.$$

Negative Binomial Distribution

$$f(x) = \frac{r(r+1) \dots (r+x-1)}{x!} \frac{\beta^x}{(1+\beta)^{r+x}}$$
$$= \binom{r+x-1}{x} \frac{\beta^x}{(1+\beta)^{r+x}}.$$

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

$$f(1) = \frac{r\beta}{(1+\beta)^{r+1}}.$$

$$f(2) = \frac{r(r+1)}{2} \frac{\beta^2}{(1+\beta)^{r+2}}.$$

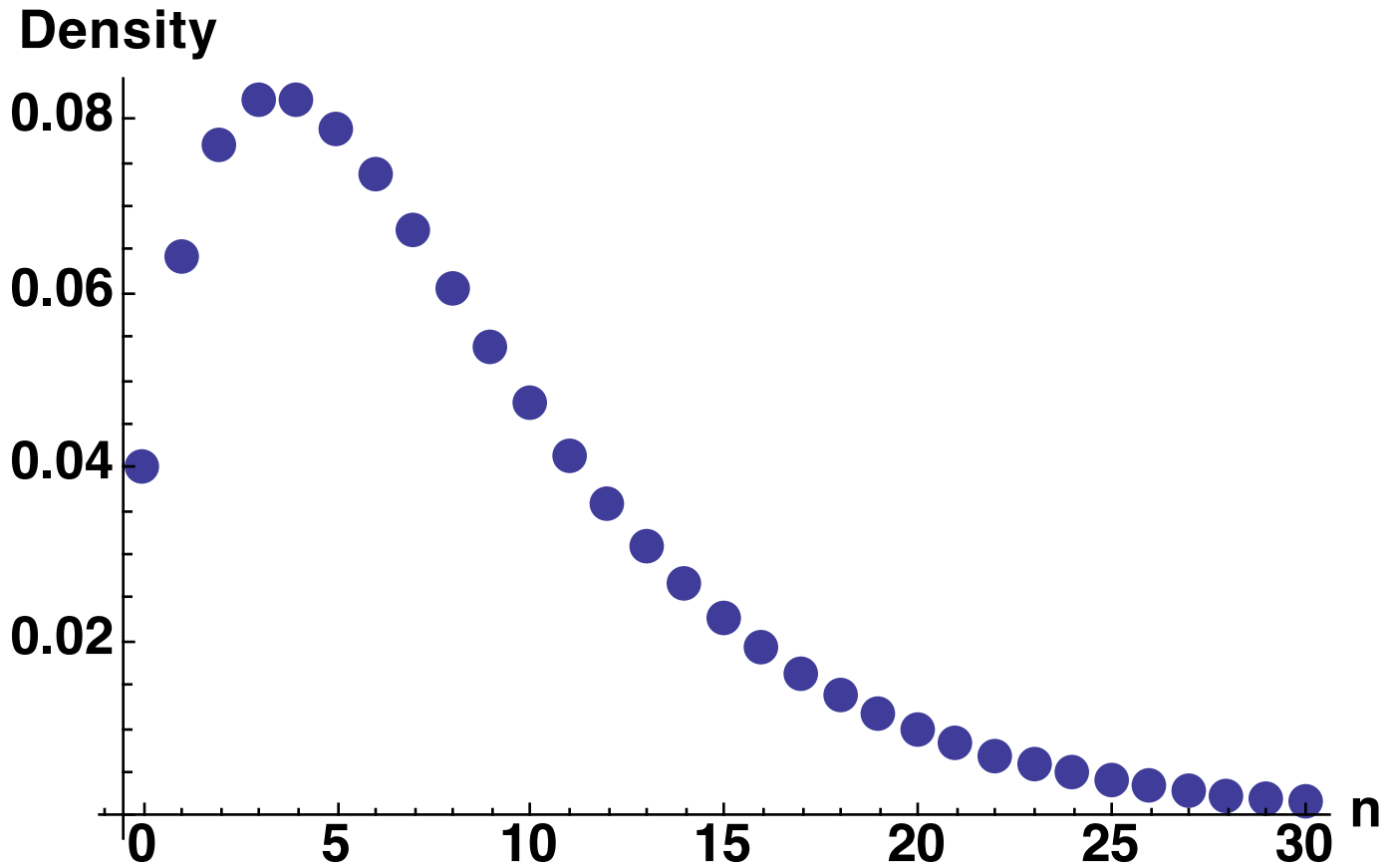
$$f(3) = \frac{r(r+1)(r+2)}{6} \frac{\beta^3}{(1+\beta)^{r+3}}$$

Negative Binomial Distribution

$$\text{Mean} = r \beta \quad \text{Variance} = r \beta(1+\beta) > \text{Mean.}$$

Negative Binomial Distrib. with $r = 2$ and $\beta = 4$,

$$f(x) = \frac{(2)(3)\dots(x+1)}{x!} \frac{4^x}{5^{x+2}} = (x+1) (0.04) (0.8^x):$$



Negative Binomial with $r = 1$ is a Geometric.

For the Negative Binomial Distribution with parameters β and r , with r integer \Leftrightarrow the sum of r independent Geometric distributions each with mean β .

Neg Bin. \Leftrightarrow Gamma Geometric \Leftrightarrow Exponential

The sum of independent Negative Binomial Distributions with the same β , is Negative Binomial with the sum of the r parameters.

6.33. CAS3, 11/06, Q.23

An actuary has determined that the number of claims follows a negative binomial distribution with mean 3 and variance 12.

Calculate the probability that the number of claims is at least 3 but less than 6.

CAS3, 11/06, Q.23. B.

$$r\beta = 3. \quad r\beta(1+\beta) = 12.$$

$$\Rightarrow 1 + \beta = 12/3 = 4. \Rightarrow \beta = 3. \Rightarrow r = 1.$$

$$f(3) + f(4) + f(5) = \frac{3^3}{4^4} + \frac{3^4}{4^5} + \frac{3^5}{4^6} = \mathbf{0.244}.$$

Comment: We have fit via Method of Moments.
Since $r = 1$, this is a Geometric Distribution.

Normal Distribution

Standard Normal, with mean zero and standard deviation one, which is in the table attached to the exam: Distribution Function $\Phi[x]$

As shown at the very beginning of the Tables:

Density function $\phi[x] = \frac{\exp[-x^2 / 2]}{\sqrt{2\pi}}$, $-\infty < x < \infty$.

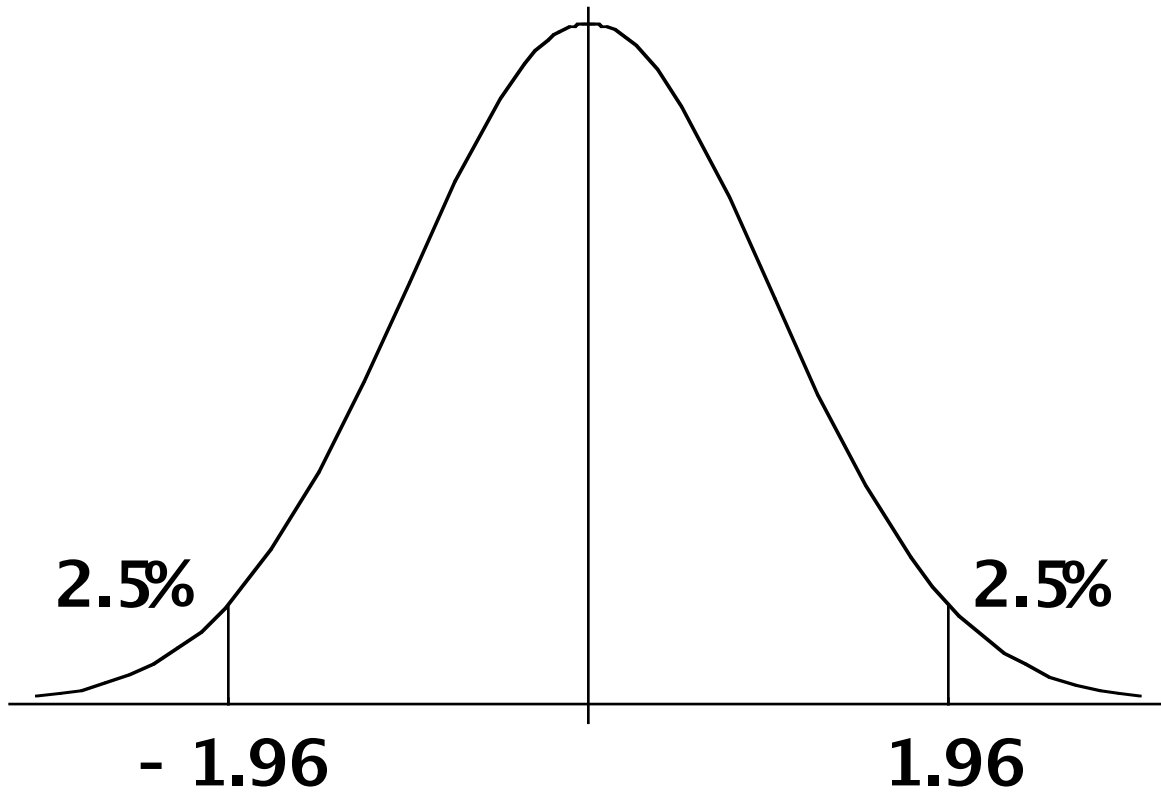
$$F(x) = \Phi\left[\frac{x - \mu}{\sigma}\right]$$

$$f(x) = \phi\left[\frac{x - \mu}{\sigma}\right] / \sigma = \frac{\exp\left[-\frac{(x - \mu)^2}{2\sigma^2}\right]}{\sigma\sqrt{2\pi}}.$$

$$\text{Mean} = \mu \quad \text{Variance} = \sigma^2$$

Skewness = 0 (distribution is symmetric)

Standard Normal Dist, 95% Confidence Interval



NORMAL DISTRIBUTION TABLE

(Standard Normal with $\mu = 0$ and $\sigma = 1$)

x	0	1	2	3
0.0	0.5000	0.8413	0.9772	0.9987
0.1	0.5398	0.8643	0.9821	0.9990
0.2	0.5793	0.8849	0.9861	0.9993
0.3	0.6179	0.9032	0.9893	0.9995
0.4	0.6554	0.9192	0.9918	0.9997
0.5	0.6915	0.9332	0.9938	0.9998
0.6	0.7257	0.9452	0.9953	0.9998
0.7	0.7580	0.9554	0.9965	0.9999
0.8	0.7881	0.9641	0.9974	0.9999
0.9	0.8159	0.9713	0.9981	1.0000

$$\Phi(z) = \Pr(Z < z)$$

$$z$$

0.800

0.842

0.850

1.036

0.900

1.282

0.950**1.645****0.975****1.960**

0.990

2.326

0.995

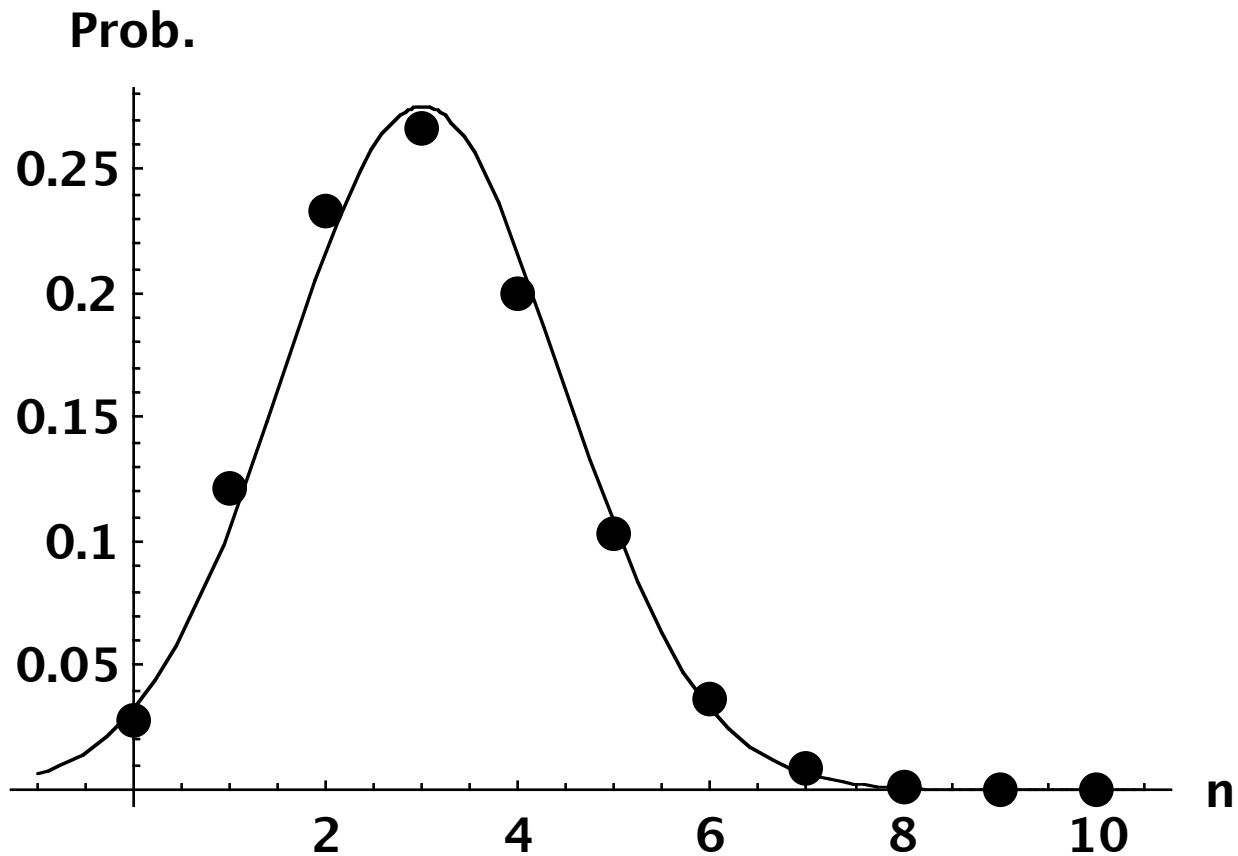
2.576

When using the normal distribution, choose the nearest z-value to find the probability, or if the probability is given, chose the nearest z-value. No interpolation should be used.

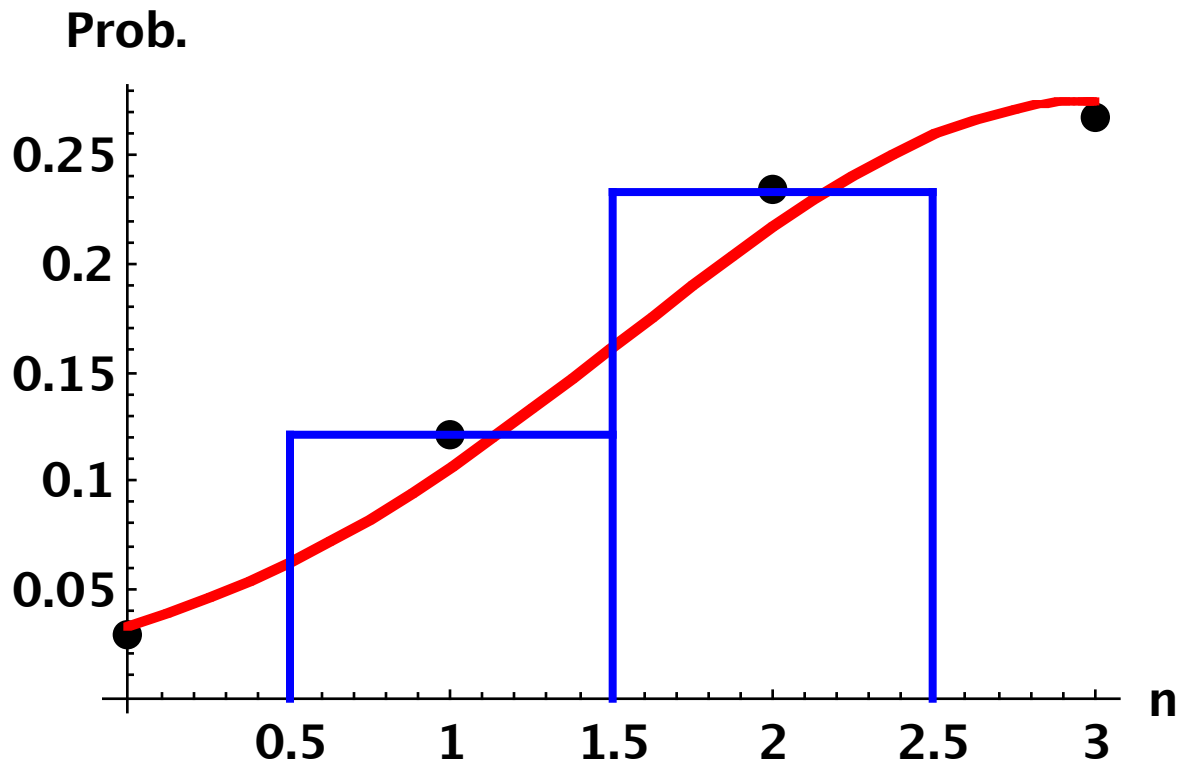
Page 125.

A Binomial Distribution with $m = 10$ and $q = 0.3$,
has mean: $(10)(0.3) = 3$,
and variance: $(10)(0.3)(0.7) = 2.1$.

A graph of this Binomial and the Normal
Distribution with the **same mean and variance**:



A Binomial Distribution with $m = 10$ and $q = 0.3$.
 The exact probability of one or two claims is $f(1) + f(2)$, the sum of the area of two rectangles:



The rectangles go from $1 - 1/2 = 0.5$
 to $2 + 1/2 = 2.5$.

The Normal Approximation is the area under the Normal Distribution with the same mean and variance, from 0.5 to 2.5.

This **continuity correction** is used with the Normal Approximation for discrete variables.

More than 15 claims \Leftrightarrow At least 16 claims

15 15.5 16

$\downarrow \rightarrow$

$$\text{Prob}[\text{More than 15 claims}] \cong 1 - \Phi\left[\frac{15.5 - \mu}{\sigma}\right].$$

For a frequency distribution with mean 14 and standard deviation 2:

$$\text{Prob}[\text{At least 16 claims}] = \text{Prob}[> 15 \text{ claims}] \cong$$

$$1 - \Phi\left[\frac{15.5 - \mu}{\sigma}\right] = 1 - \Phi\left[\frac{15.5 - 14}{2}\right] =$$

$$1 - \Phi[0.75] = 1 - 0.7734 = 22.66\%.$$

Page 127.

μ = mean of frequency distribution

σ = standard deviation of frequency distrib.

Prob[$i \leq \# \text{ claims} \leq j$] \cong

$$\Phi\left[\frac{(j + 0.5) - \mu}{\sigma}\right] - \Phi\left[\frac{(i - 0.5) - \mu}{\sigma}\right].$$

Less than 12 claims \Leftrightarrow At most 11 claims

\Leftrightarrow 11 claims or less.

11 11.5 12

\leftarrow

$$\text{Prob}[\text{Less than 12 claims}] \cong \Phi\left[\frac{11.5 - \mu}{\sigma}\right].$$

7.13. You are given the following:

- Sue takes an actuarial exam with 40 multiple choice questions, each of equal value.
- Sue knows the answers to 13 questions and answers them correctly.
- Sue guesses at random on the remaining 27 questions, with a $1/5$ chance of getting each such question correct, with each question independent of the others.

If 22 correct answers are needed to pass the exam, what is the probability that Sue passed her exam? Use the Normal Approximation.

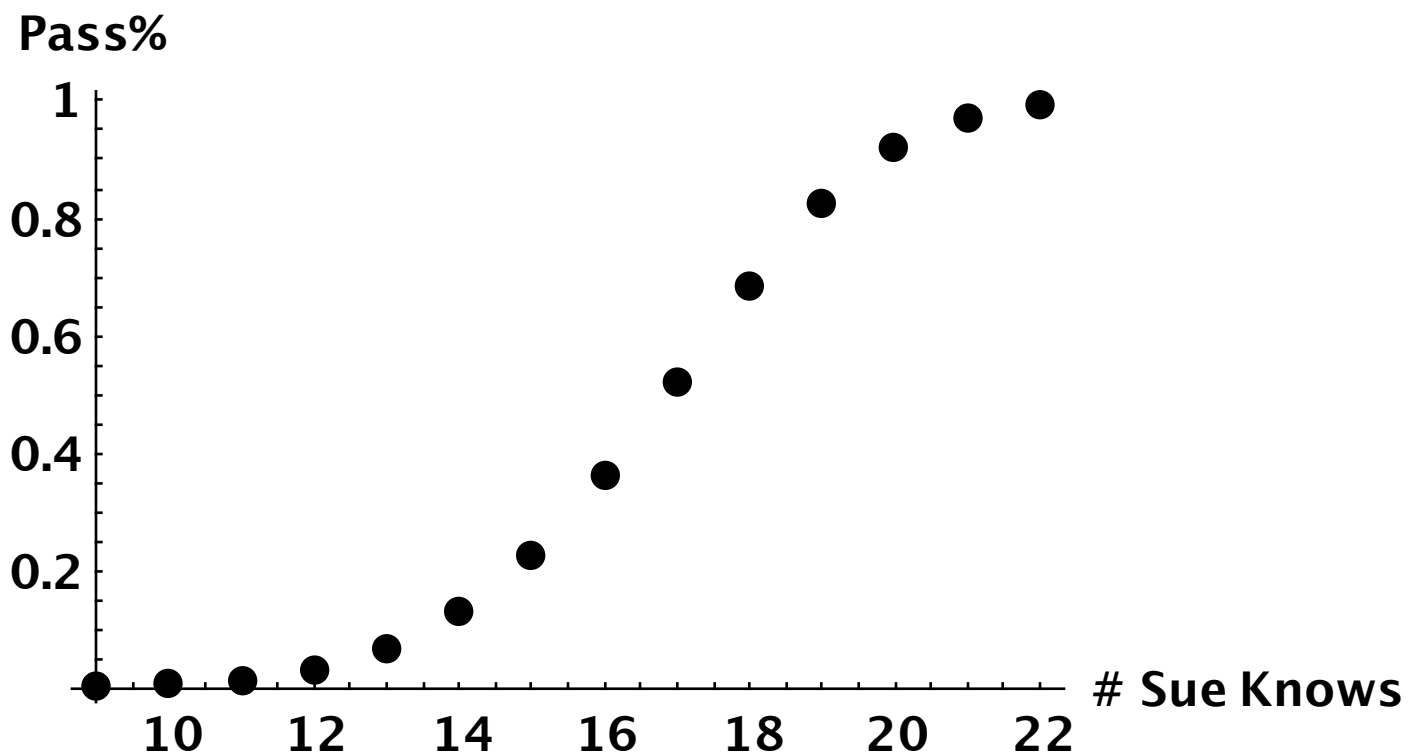
- A. 4% B. 5% C. 6% D. 7% E. 8%

7.13. D. The number of correct guesses is Binomial with parameters: $m = 27$ and $q = 1/5$, with mean: $(1/5)(27) = 5.4$ and variance: $(1/5)(4/5)(27) = 4.32$.

Therefore, $\text{Prob}(\# \text{ correct guesses} \geq 9) \cong$

$$1 - \Phi\left[\frac{8.5 - 5.4}{\sqrt{4.32}}\right] = 1 - \Phi[1.49] = \mathbf{6.8\%}.$$

Here is a graph of Sue's chance of passing, given that she knows the answers to a certain number of questions and guesses on the rest:

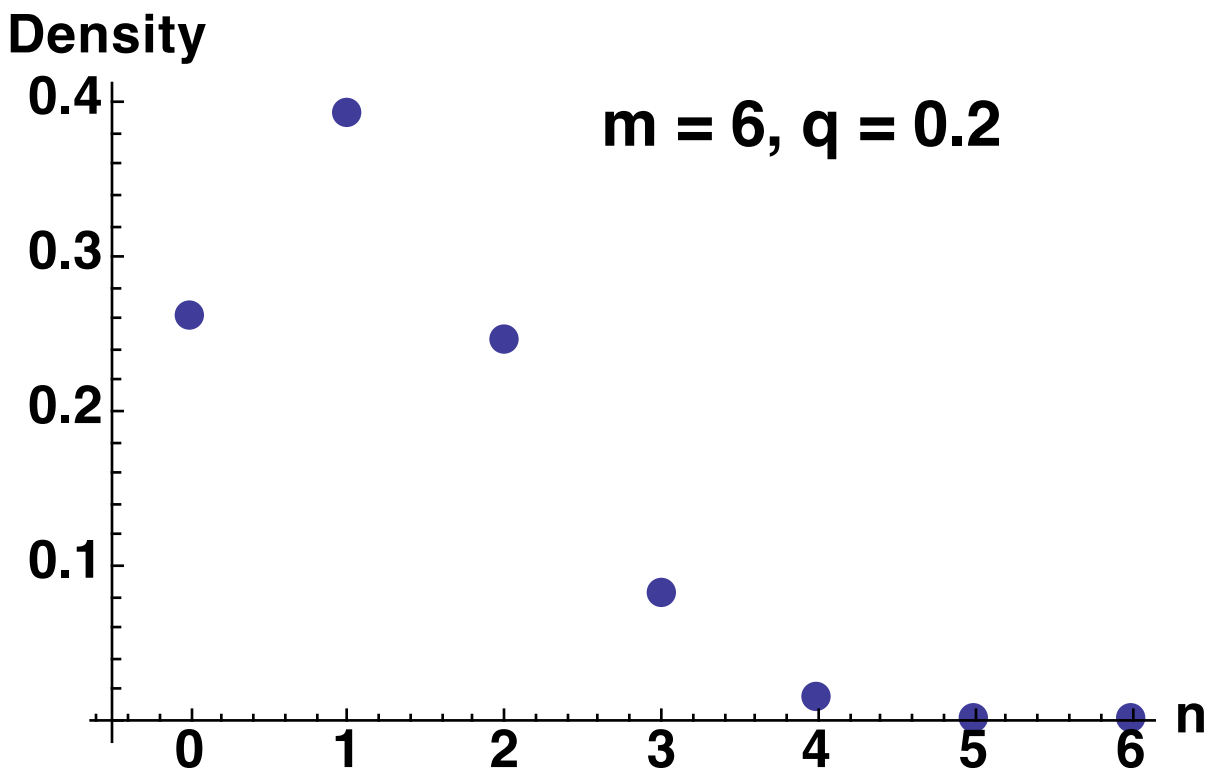


$$\begin{aligned} \text{Skewness} &= \frac{E[(X - E[X])^3]}{\text{STDDEV}^3} \\ &= \frac{\text{Third Central Moment}}{\text{Variance}^{1.5}}. \end{aligned}$$

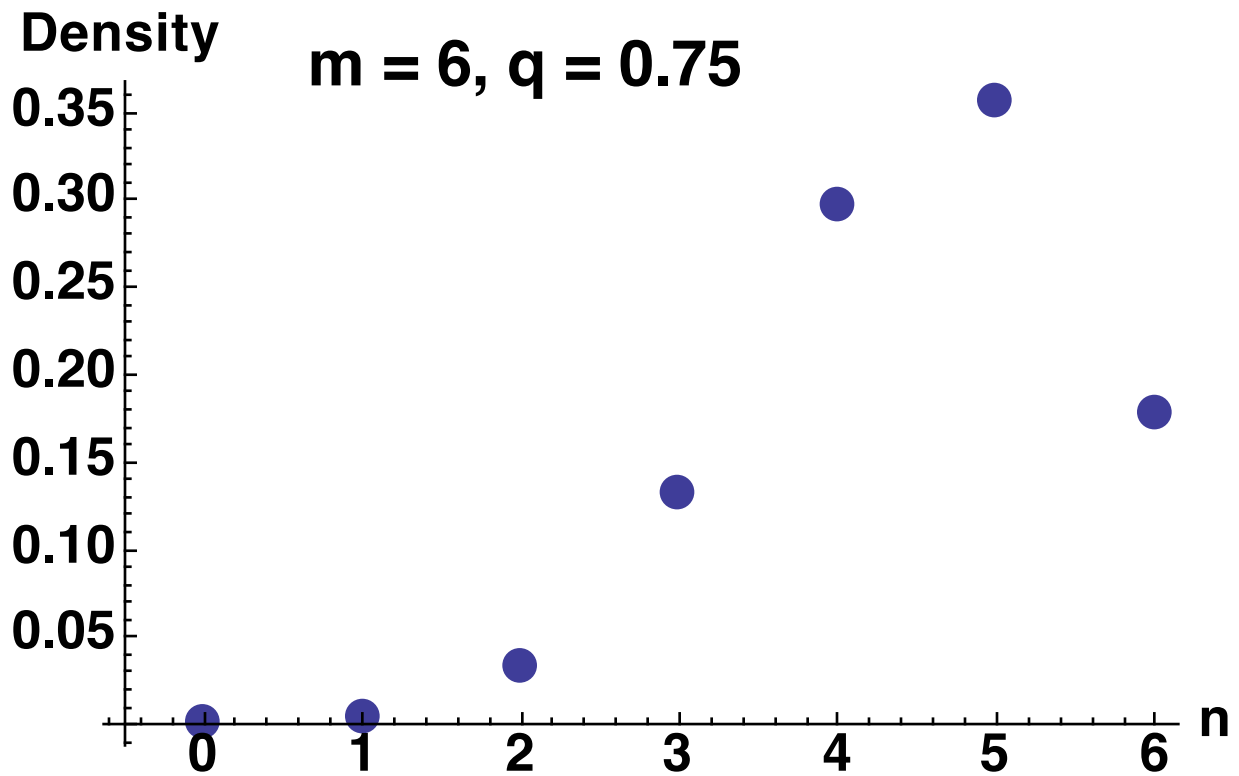
Positive skewness: Poisson & Negative Binomial.

As $q \rightarrow 0$, Binomial \rightarrow Poisson.

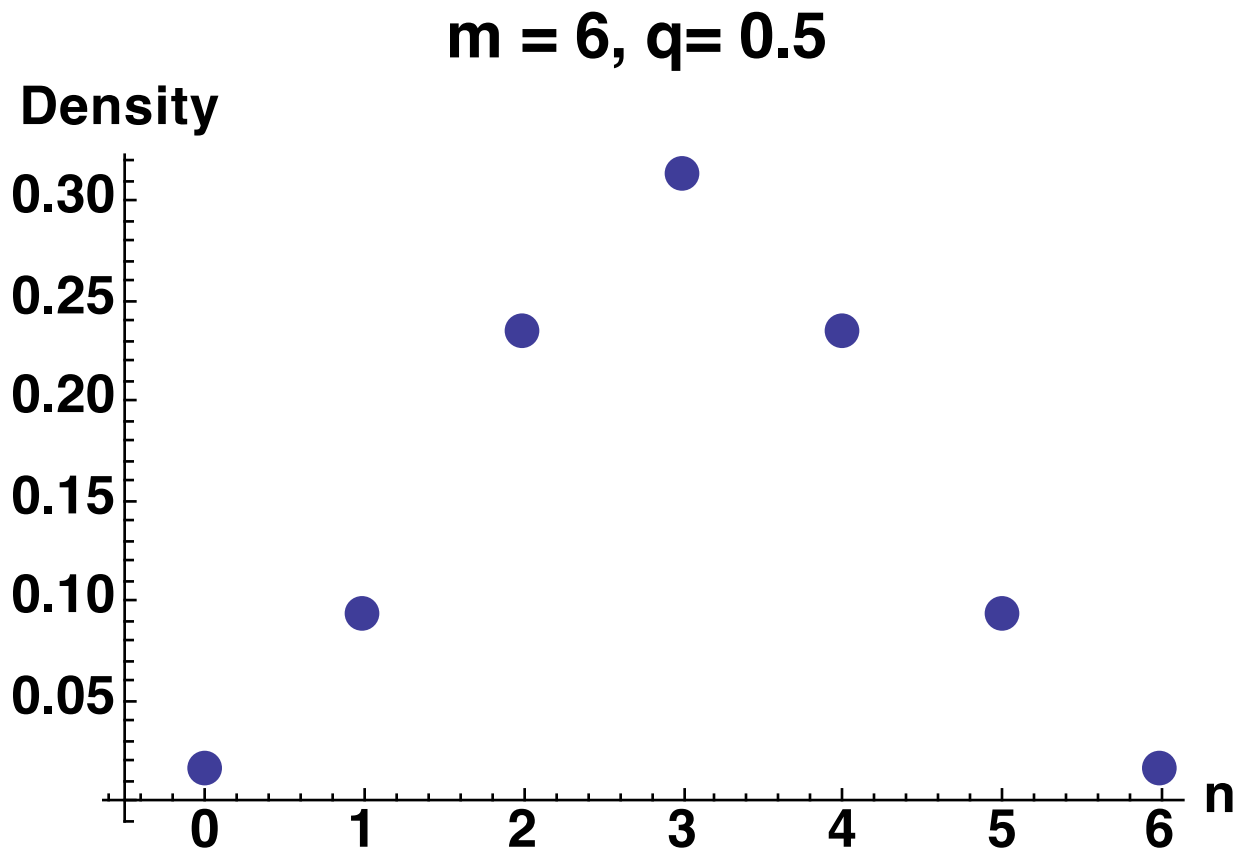
Binomial with $q < 1/2$, **positive** skewness.



Binomial with $q > 1/2$, **negative** skewness.



Binomial with $q = 1/2$ is symmetric: 0 skewness.



Probability Generating Function, p.g.f.:

$$\infty$$

$$P(z) = \text{Expected Value of } z^n \equiv \sum_{n=0}^{\infty} f(n) z^n.$$

Poisson: $P(z) = \text{Exp}[\lambda(z - 1)]$.

p.g.f. of the sum of independent frequencies is the product of the individual p.g.f.s.

The distribution determines the probability generating function and vice versa.

$$f(n) = \left(\frac{d^n P(z)}{dz^n} \right)_{z=0} / n!.$$

$$f(0) = P(0).$$

$$f(1) = P'(0).$$

$$f(2) = P''(0) / 2.$$

Factorial Moments

$$\begin{aligned} \text{nth factorial moment} &= \mu_{(n)} \\ &= E[X(X-1) \dots (X + 1 - n)]. \end{aligned}$$

$$\mu_{(2)} = E[X(X-1)] = E[X^2] - E[X].$$

$$\text{Poisson: } \mu_{(n)} = \lambda^n.$$

$$\mu_{(n)} = \left(\frac{d^n P(z)}{dz^n} \right)_{z=1}.$$

$$E[X] = P'(1).$$

(a, b, 0) Class of Distributions:

<u>Distribution</u>	<u>Mean</u>	<u>Variance</u>
Binomial, Var. < Mean	$m q$	$m q (1 - q)$
Poisson, Var. = Mean	λ	λ
Negative Binomial Variance > Mean	$r \beta$	$r \beta (1 + \beta)$

$$\frac{f(x+1)}{f(x)} = a + \frac{b}{x+1}, \quad x = 0, 1, 2, \dots$$

$$f(x+1) = \left\{ a + \frac{b}{x+1} \right\} f(x), \quad x = 0, 1, 2, \dots$$

$$f(1) = (a + b) f(0). \quad f(2) = (a + b/2) f(1).$$

$$f(3) = (a + b/3) f(2). \quad f(4) = (a + b/4) f(3).$$

$$\frac{p_k}{p_{k-1}} = a + \frac{b}{k}, \quad k = 1, 2, 3 \dots$$

a and b are in the Tables attached to the exam:

<u>Distribution</u>	<u>a</u>	<u>b</u>
Binomial	$-q / (1 - q)$	$(m + 1)q / (1 - q)$
Poisson	0	λ
Negative Binomial	$\beta / (1 + \beta)$	$(r - 1)\beta / (1 + \beta)$

11.19. 3, 11/02, Q.28 & 2009 Sample Q.94

X is a discrete random variable with a probability function which is a member of the $(a,b,0)$ class of distributions.

You are given:

(i) $P(X = 0) = P(X = 1) = 0.25$

(ii) $P(X = 2) = 0.1875$

Calculate $P(X = 3)$.

- (A) 0.120 (B) 0.125 (C) 0.130 (D) 0.135 (E) 0.140

3, 11/02, Q.28. B. For a member of the $(a,b,0)$ class of distributions, $f(x+1) / f(x) = a + \{b / (x+1)\}$.

$$f(1)/f(0) = a + b. \Rightarrow 0.25/0.25 = 1 = a + b.$$

$$f(2)/f(1) = a + b/2. \Rightarrow 0.1875/0.25 = 0.75 = a + b/2.$$

Therefore, $a = 0.5$ and $b = 0.5$.

$$f(3) = f(2) (a + b/3) = (0.1875) (0.5 + 0.5/3) = \mathbf{0.125}.$$

Alternately, once one solves for a and b ,

$a > 0 \Rightarrow$ Negative Binomial Distribution.

$$1/2 = a = \beta / (1 + \beta). \Rightarrow \beta = 1.$$

$$1/2 = b = (r - 1) \beta / (1 + \beta). \Rightarrow r - 1 = 1. \Rightarrow r = 2.$$

$$f(3) = \frac{r(r + 1)(r + 2)}{3!} \frac{\beta^3}{(1+\beta)^{r+3}} = \frac{(2)(3)(4)}{6} \frac{1^3}{2^5} = \mathbf{1/8}.$$

Accident Profiles

For a member of the $(a, b, 0)$ class:

$$\frac{f(x+1)}{f(x)} = a + \frac{b}{x+1}$$

$$\Leftrightarrow (x+1) f(x+1) / f(x) = (x+1) a + b = ax + a + b.$$

It is a straight line with slope a , and intercept $a + b$.

Thus graphing $\frac{(x+1) f(x+1)}{f(x)}$ can be a useful

method of determining whether one of these three distributions fits the given data.

Thus graphing $\frac{(x+1) f(x+1)}{f(x)}$ can be a useful

method of determining whether one of these three distributions fits the given data.

If a straight line does seem to fit this “accident profile”, then one should use a member of the $(a, b, 0)$ class.

The slope determines which of the three distributions is likely to fit:

if the slope is close to zero then a Poisson,
if significantly negative then a Binomial, and
if significantly positive then a Negative Binomial.

12.5. You are given the following distribution of the number of claims per policy during a one-year period for 20,000 policies.

<u>Number of claims per policy</u>	<u>Number of Policies</u>
0	6503
1	8199
2	4094
3	1073
4	128
5	3
6+	0

Which of the following distributions would be the most appropriate model for this data?

12.5. Calculate $\frac{(x+1) f(x+1)}{f(x)} =$

$\frac{(x+1) \text{ (number of policies with } x+1 \text{ claims)}}{\text{(number of policies with } x \text{ claims)}}$.

Number of claims	Observed	$(x+1) f(x+1) / f(x)$	Differences
0	6,503	1.261	
1	8,199	0.999	-0.262
2	4,094	0.786	-0.212
3	1,073	0.477	-0.309
4	128	0.117	-0.360
5	3		

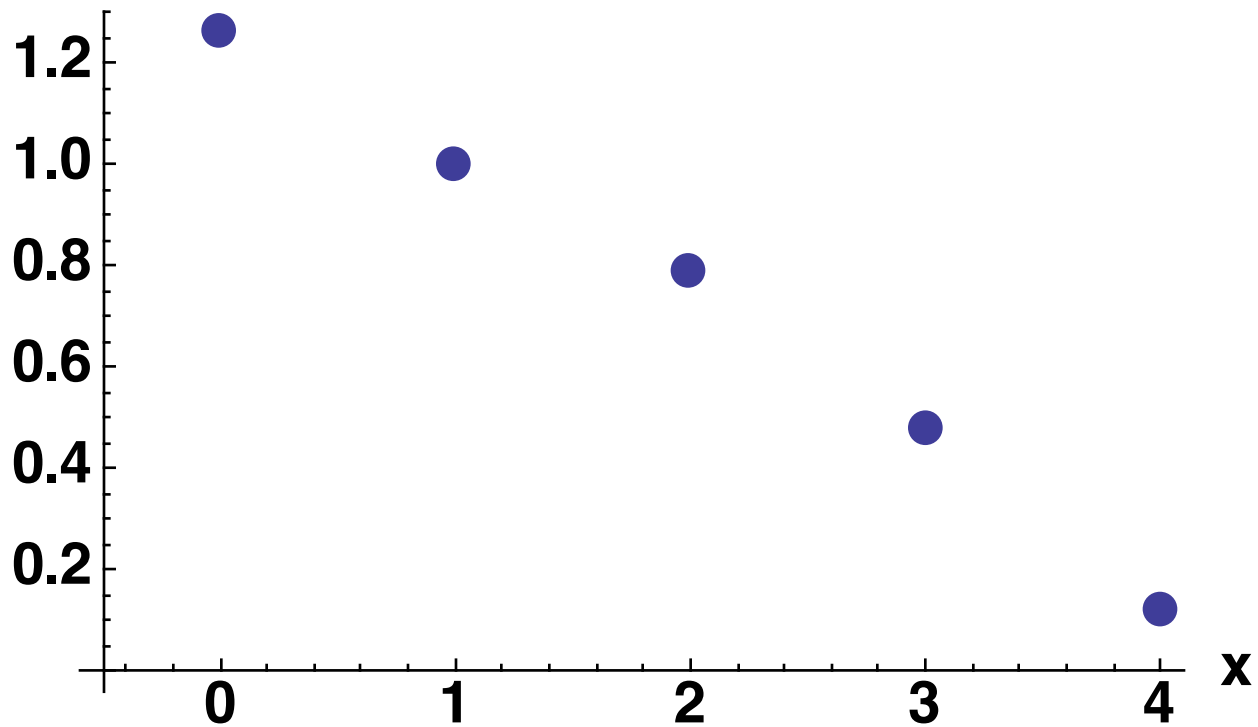
Since $\frac{(x+1) f(x+1)}{f(x)}$ is approximately linear,

we probably have a member of the $(a, b, 0)$ class.

$a = \text{slope} < 0. \Rightarrow$ Binomial Distribution.

A graph of

$$\frac{(x+1) \text{ (number of policies with } x+1 \text{ claims)}}{\text{(number of policies with } x \text{ claims)}}$$



The sample mean is 1.00665.

The sample variance is 0.804.

Since the sample mean is greater than the sample variance by a significant amount, **if** this is a member of the $(a, b, 0)$ class then it is a Binomial Distribution.

Zero-Truncated Distributions

If f is a distribution on $0, 1, 2, 3, \dots$, then

$$p_k^T = \frac{f(k)}{1 - f(0)} \text{ is a distribution on } 1, 2, 3, \dots$$

Loss Models uses the notation p_k^T for the densities of a zero-truncated distribution.

The moments of a zero-truncated distribution are given in terms of those of the corresponding untruncated distribution, f , by:

$$E^{\text{Truncated}}[X^n] = \frac{E_f[X^n]}{1 - f(0)}.$$

13.4. The number of vehicles involved in an automobile accident is given by a Zero-Truncated Binomial Distribution with parameters $q = 0.3$ and $m = 5$.

What is the chance of observing exactly 3 vehicles involved in an accident?

13.4. D. For a non-truncated Binomial,

$$f(3) = \frac{5!}{3! 2!} 0.3^3 0.7^2 = 0.1323.$$

For the zero-truncated distribution one gets the density by dividing by $1 - f(0)$:

$$0.1323 / (1 - 0.7^5) = 15.9\%.$$

Page 231

The Logarithmic Distribution with parameter β has support equal to the positive integers:

$$f(x) = \frac{\beta^x}{(1 + \beta)^x} \frac{1}{x \ln(1 + \beta)} \quad \text{for } x = 1, 2, 3, \dots$$

Page 232. For the **(a, b, 1) class of frequency distributions:**

$$\frac{f(x+1)}{f(x)} = a + \frac{b}{x + 1}, \text{ for } x \geq 1.$$

For the (a, b, 0) class of frequency distributions we required that the same relationship hold, but starting instead at $x = 0$.

The (a, b, 1) class consists of:
the (a, b, 0) class,
the zero-truncated distributions,
the logarithmic distribution,
and the zero-modified distributions.

For a Zero-Modified distribution, an arbitrary amount of probability has been placed at zero. The remaining probability has been spread on $x = 1, 2, 3,$ proportional to some density $f(x)$ on $x = 0, 1, 2, 3,$

Density at zero is p_0^M ,

$$p_k^M = f(x) \frac{1 - p_0^M}{1 - f(0)}, x = 1, 2, 3, \dots$$

is a distribution on $0, 1, 2, 3, \dots$

Loss Models uses the notation \mathbf{p}_k^M for the densities of a zero-modified distribution.

As shown in Appendix B:

$$p_k^M = (1 - p_0^M) p_k^T = (1 - p_0^M) p_k / (1 - p_0).$$

The moments of a zero-modified distribution are given in terms of those of the corresponding unmodified distribution, f , by:

$$E_{\text{Modified}}[X^n] = E_f[X^n] \frac{1 - p_0^M}{1 - f(0)}.$$

The probability generating functions of the **zero-truncated** distributions are given in Appendix B.

The probability generating functions of the **zero-modified** distributions are given in terms of those of the corresponding **zero-truncated** distribution, by:

$$P^M[z] = p_0^M + (1 - p_0^M) P^T[z].$$

14.20. The number of claims per year is given by a Zero-Modified Poisson Distribution with parameter $\lambda = 2.5$, and with 30% chance of zero claims. What is the chance of observing 2 claims over the coming year?

14.20. B. For an unmodified Poisson with $\lambda = 2.5$:

$$f(0) = e^{-2.5}.$$

$$f(2) = (2.5^2) e^{-2.5} / 2! = 0.2565.$$

For the zero-modified distribution one gets the densities for values other than zero by multiplying by $(1 - 0.3)$ and dividing by $1 - e^{-2.5}$:

$$(0.2565)(1 - 30\%) / (1 - e^{-2.5}) = 19.56\%.$$

Compound Frequency Distributions

Compound distributions are mathematically equivalent to aggregate distributions.

The formula for their variance will be covered with aggregate distributions.

Mixed Frequency Distributions

Each insured has a Poisson frequency, but there are four types:

<u>Type</u>	<u>Lambda</u>	<u>Probability</u>
Excellent	1	40%
Good	2	30%
Bad	3	20%
Ugly	4	10%

For example, for an Ugly insured the chance of 6 claims is: $\frac{4^6 e^{-4}}{6!} = 10.4\%$.

For an insured picked at random of unknown type, the probability of 6 claims is:

$$(.4)(.05\%) + (.3)(1.20\%) + (.2)(5.04\%) + (.1)(10.42\%) = 2.43\%.$$

The density function of the mixture, is the mixture of the density functions.

<u>Type</u>	<u>Lambda</u>	<u>Probability</u>
Excellent	1	40%
Good	2	30%
Bad	3	20%
Ugly	4	10%

Number of Claims	Prob. Excellent	Prob. Good	Prob. Bad	Prob. Ugly	Prob. All
0	0.3679	0.1353	0.0498	0.0183	0.1995
1	0.3679	0.2707	0.1494	0.0733	0.2656
2	0.1839	0.2707	0.2240	0.1465	0.2142
3	0.0613	0.1804	0.2240	0.1954	0.1430
4	0.0153	0.0902	0.1680	0.1954	0.0863
5	0.0031	0.0361	0.1008	0.1563	0.0478
6	0.0005	0.0120	0.0504	0.1042	0.0243
7	0.0001	0.0034	0.0216	0.0595	0.0113
8	0.0000	0.0009	0.0081	0.0298	0.0049
9	0.0000	0.0002	0.0027	0.0132	0.0019
10	0.0000	0.0000	0.0008	0.0053	0.0007
SUM	1.0000	1.0000	1.0000	1.0000	1.0000

For 6 claims: $(40\%)(.05\%) + (30\%)(1.20\%)$
 $+ (20\%)(5.04\%) + (10\%)(10.42\%) = 2.43\%$.

This is an example of a discrete mixture.
 There are also continuous mixtures.

17.14. Each insured's claim frequency follows a Negative Binomial Distribution, with $r = 0.8$.

There are two types of insureds as follows:

<u>Type</u>	<u>A Priori Probability</u>	<u>β</u>
-------------	-----------------------------	---------------------------

A	70%	0.2
---	-----	-----

B	30%	0.5
---	-----	-----

What is the chance of an insured picked at random having 1 claim next year?

A. 13% B. 14% C. 15% D. 16% E. 17%

$$\mathbf{17.14. B.} \quad f(1) = \frac{r \beta}{(1 + \beta)^{r+1}}.$$

$$\text{For Type A: } f(1) = \frac{(0.8) (0.2)}{1.2^{1.8}} = \mathbf{11.52\%}.$$

$$\text{For Type B: } f(1) = \frac{(0.8) (0.5)}{1.5^{1.8}} = \mathbf{19.28\%}.$$

$$(70\%) (\mathbf{11.52\%}) + (30\%) (\mathbf{19.28\%}) = \mathbf{13.85\%}.$$

17.21. For a given value of q , the number of claims is Binomial distributed with parameters $m = 3$ and q .

In turn q is distributed uniformly from 0 to 0.4.

What is the chance that zero claims are observed?

17.21. D. Given q , we have a Binomial with parameters $m = 3$ and q .

The chance that we observe 0 claims is: $(1 - q)^3$.

$$\pi(q) = 1 / (0.4 - 0) = 2.5, \text{ for } 0 \leq q \leq 0.4.$$

$$f(0) = \int_0^{0.4} f(0 | q) \pi(q) dq = \int_0^{0.4} (1 - q)^3 (2.5) dq$$

$$q = 0.4$$

$$= (-2.5/4) (1 - q)^4 \Big|_{q=0}^{q=0.4} = (-0.625) (0.6^4 - 1^4) = \mathbf{0.544}.$$

Complete Gamma Function:

$$\Gamma(\alpha) = \int_0^{\infty} t^{\alpha-1} e^{-t} dt = \theta^{-\alpha} \int_0^{\infty} t^{\alpha-1} e^{-t/\theta} dt .$$

For α integer: $\Gamma(\alpha) = (\alpha - 1) !$ $\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)$.

$\Gamma(1) = 1$. $\Gamma(2) = 1$. $\Gamma(3) = 2$. $\Gamma(4) = 6$. $\Gamma(5) = 24$.

Incomplete Gamma Function:

$$\Gamma(\alpha; x) = \int_0^x t^{\alpha-1} e^{-t} dt / \Gamma(\alpha).$$

$$\Gamma(\alpha; 0) = 0. \quad \Gamma(\alpha; \infty) = 1.$$

Gamma Distribution: $F(x) = \Gamma(\alpha; x/\theta)$.

Gamma Distribution

$F(x) = \Gamma(\alpha; x/\theta)$. Two parameters α and θ .

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

$$E[X] = \alpha \theta.$$

$$E[X^2] = \alpha(\alpha + 1) \theta^2.$$

$$\text{Var}[X] = \alpha \theta^2.$$

For the important special case $\alpha = 1$, we have an Exponential distribution:

$$f(x) = e^{-x/\theta} / \theta, x \geq 0.$$

The density of a Gamma Distrib. integrates to 1.

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

$$\int_0^{\infty} t^{\alpha-1} e^{-t/\theta} dt = \Gamma(\alpha) \theta^{\alpha}.$$

Can rewrite this with $n = \alpha - 1$, and $c = 1/\theta$:

$$\int_0^{\infty} t^n e^{-c t} dt = n! / c^{n+1}.$$

Gamma-Poisson

The number of claims a particular policyholder makes in a year is **Poisson** with mean λ .

The λ **values of the portfolio of policyholders are Gamma distributed.**

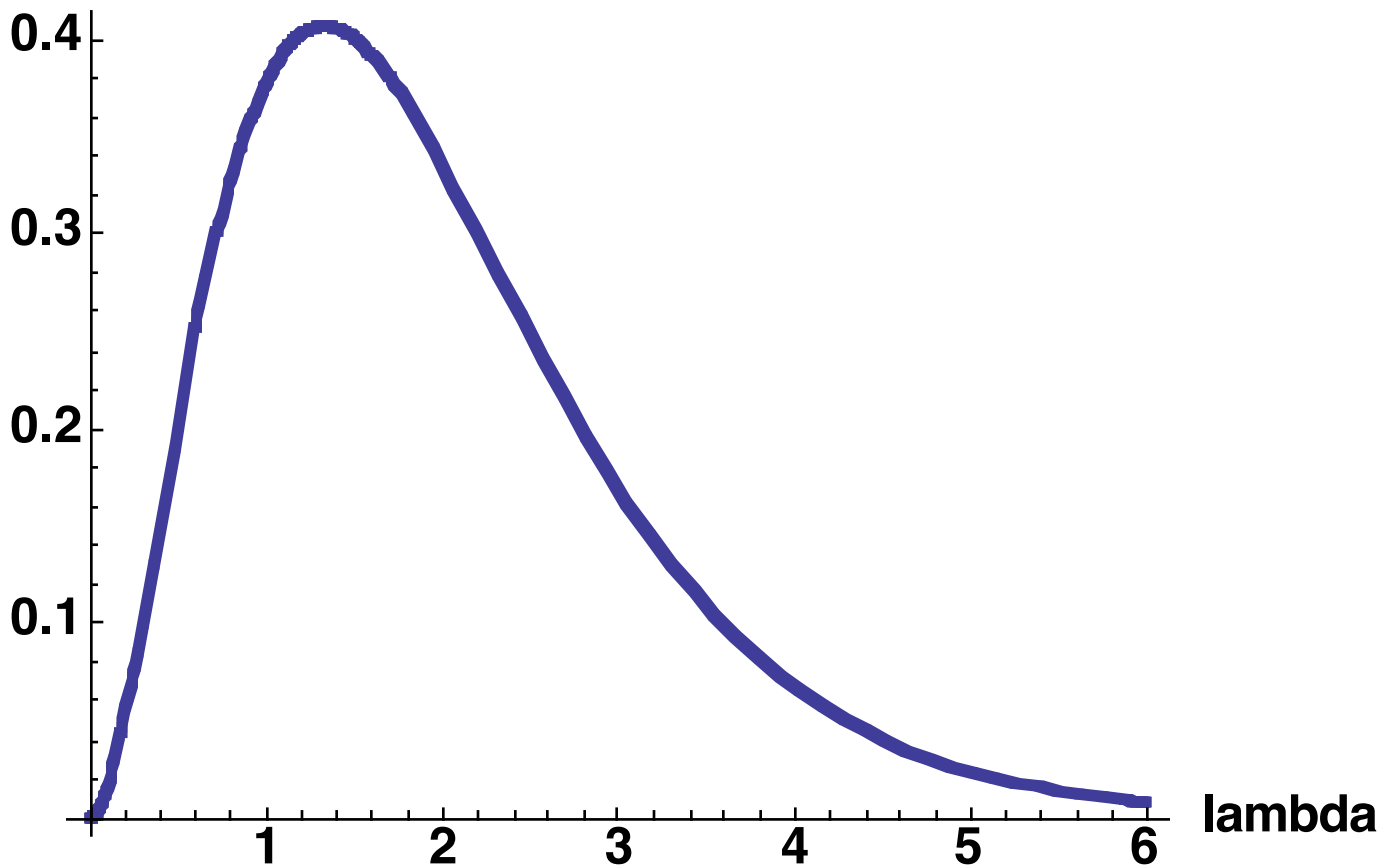
Lambda varies via a continuous distribution; this is a very important example of a continuous mixture.

Prior Distribution

The λ values of the portfolio of policyholders are Gamma distributed with $\alpha = 3$ and $\theta = 2/3$:

$$f(\lambda) = 1.6875 \lambda^2 e^{-1.5\lambda}, \lambda > 0.$$

density



Mixed Distribution

The mixed distribution is a Negative Binomial Distribution with $r = \alpha$ and $\beta = \theta$.

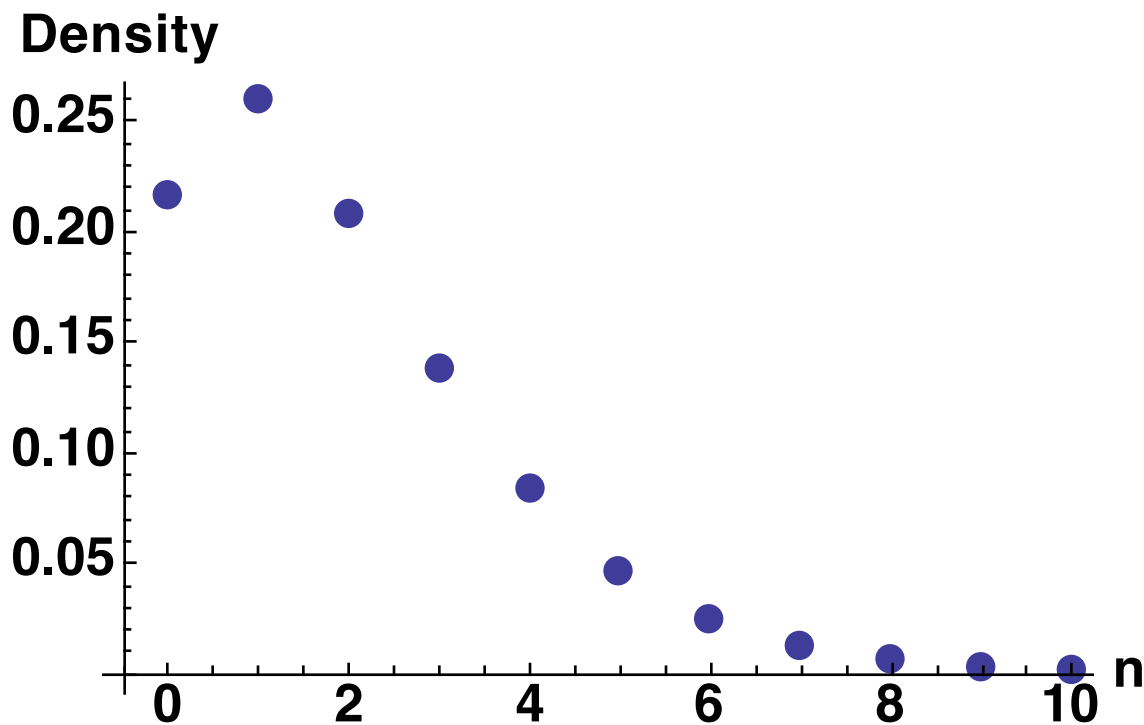
Beta rhymes with theta.

In this case, $r = \alpha = 3$, and $\beta = \theta = 2/3$.

In this case, $r = \alpha = 3$, and $\beta = \theta = 2/3$.

The chance of a policyholder chosen at random having 6 claims is:

$$\frac{\binom{3}{6} \binom{4}{5} \binom{5}{6} \binom{6}{7} \binom{8}{2/3} (2/3)^6}{6! (5/3)^{6+3}} = 2.477\%.$$



Neg. Binomial is a distribution of number of claims, while the **Gamma** is a distribution of parameters.

Overall Mean

Mean = $E[\lambda]$ = mean of the Gamma = $\alpha \theta = (3) (2/3)$

= **2**

= $(3) (2/3) = r \beta$ = mean of the Negative Binomial.

Exponential-Poisson

For the important special case $\alpha = 1$, we have an Exponential distribution of λ : $f(\lambda) = e^{-\lambda/\theta} / \theta, \lambda \geq 0$.
(See 3/11/01, Q.27.)

The mixed distribution is a Negative Binomial Distribution with $r = 1$ and $\beta = \theta$;

For the Exponential-Poisson, the mixed distribution is a Geometric Distribution with $\beta = \theta$.

We will revisit the Gamma-Poisson when we discuss Conjugate Priors.

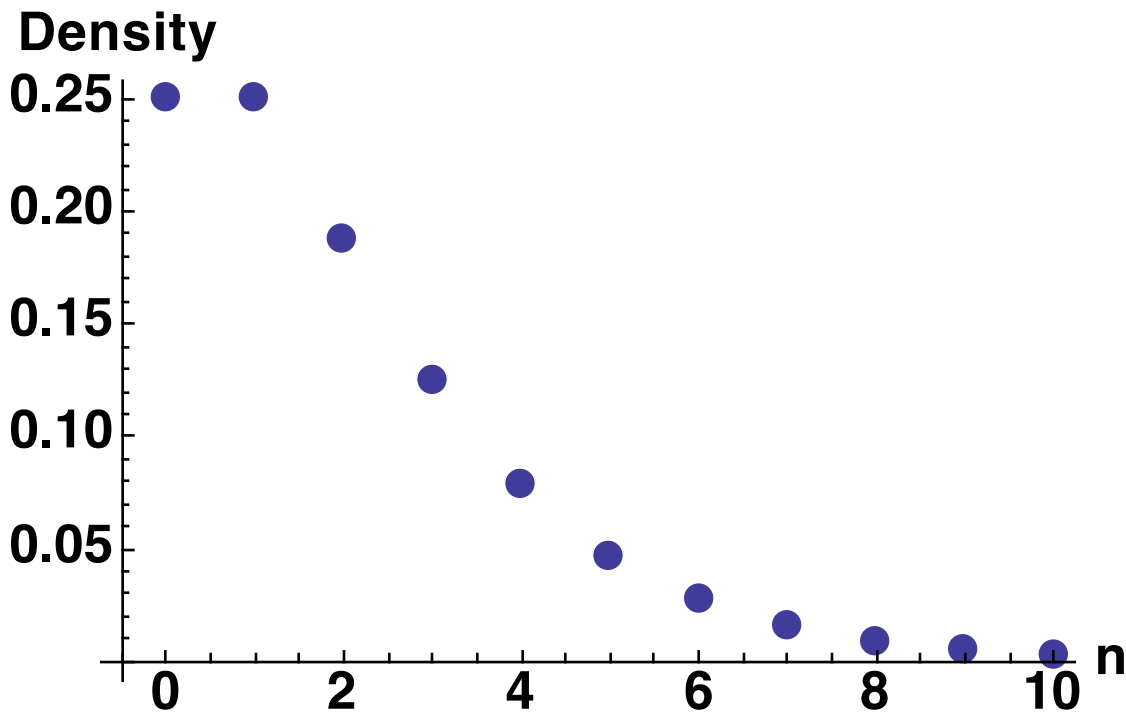
19.62. CAS3, 5/05, Q.10. Low Risk Insurance Company provides liability coverage to a population of 1,000 private passenger automobile drivers. The number of claims during a given year from this population is Poisson distributed. If a driver is selected at random from this population, his expected number of claims per year is a random variable with a Gamma distribution such that $\alpha = 2$ and $\theta = 1$. Calculate the probability that a driver selected at random will not have a claim during the year.

A. 11.1% B. 13.5% C. 25.0% D. 33.3% E. 50.0%

CAS3, 5/05, Q.10. C. Gamma-Poisson.

The mixed distribution is Negative Binomial with $r = \alpha = 2$ and $\beta = \theta = 1$.

$$f(0) = \frac{1}{(1 + \beta)^r} = \frac{1}{(1 + 1)^2} = 1/4.$$



CAS3, 5/05, Q.10. Low Risk Insurance Company provides liability coverage to a population of 1,000 private passenger automobile drivers.

The number of claims during a given year from this population is Poisson distributed.

If a driver is selected at random from this population, his expected number of claims per year is a random variable with a Gamma distribution such that $\alpha = 2$ and $\theta = 1$.

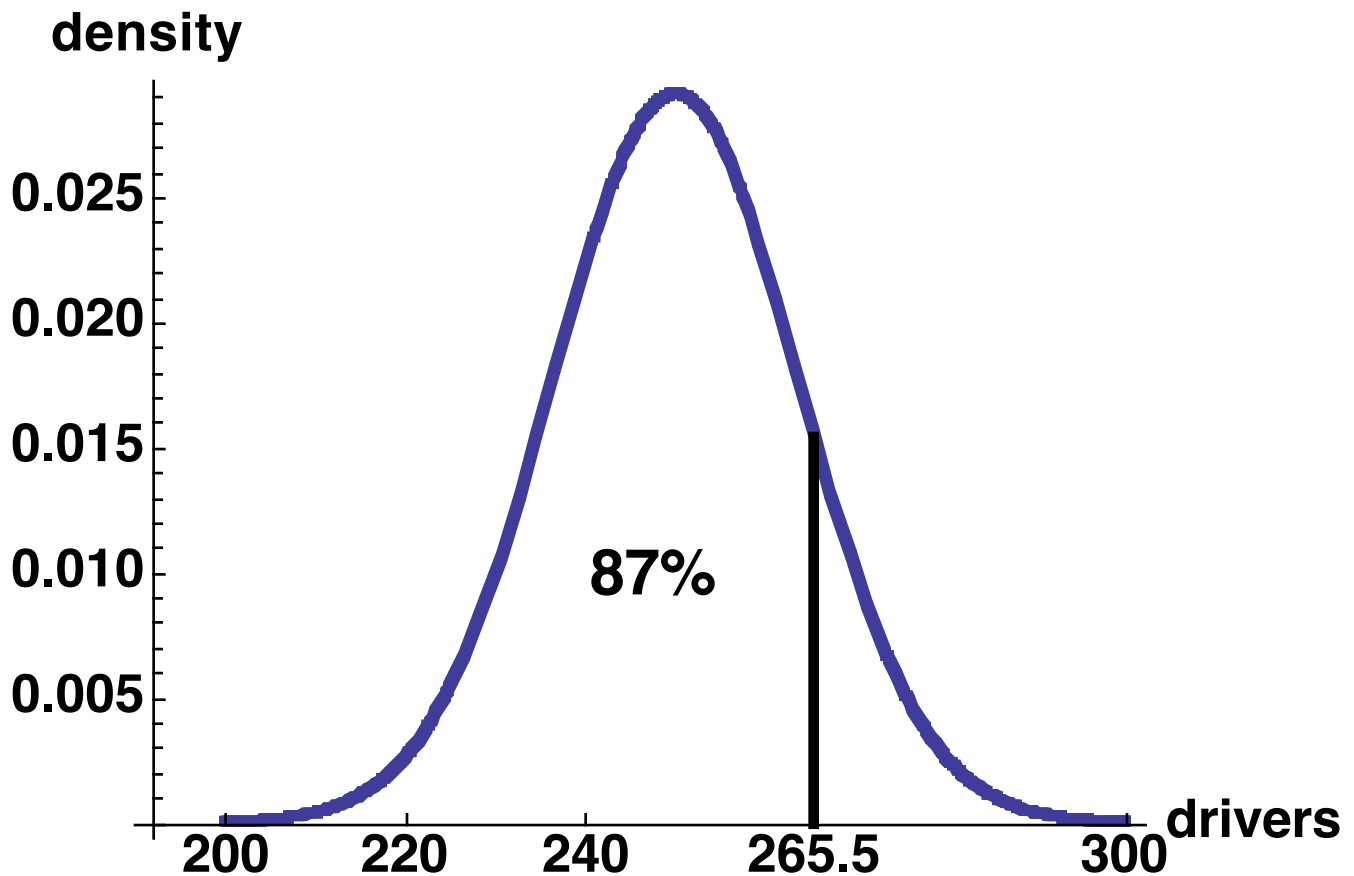
19.63. In CAS3, 5/05, Q.10, what is the probability that at most 265 of these 1000 drivers will not have a claim during the year?

19.63. E. For each driver, $\text{Prob}[\text{no claim}] = 1/4$. For 1000 independent drivers, the number of drivers with no claims is Binomial with $m = 1000$ and $q = 1/4$, with mean $m q = 250$, and variance $m q (1 - q) = 187.5$.

$\text{Prob}[\text{At most 265 claim-free drivers}] \cong$

$$\Phi\left[\frac{265.5 - 250}{\sqrt{187.5}}\right] = \Phi[1.13] = \mathbf{87.08\%}.$$

The Normal with mean 250 and variance 187.5:



CAS3, 5/05, Q.10. Low Risk Insurance Company provides liability coverage to a population of 1,000 private passenger automobile drivers.

The number of claims during a given year from this population is Poisson distributed.

If a driver is selected at random from this population, his expected number of claims per year is a random variable with a Gamma distribution such that $\alpha = 2$ and $\theta = 1$.

19.64. In CAS3, 5/05, Q.10, what is the probability that these 1000 drivers will have a total of more than 2020 claims during the year?

19.64. D.

The distribution of number of claims from a single driver is Neg. Binomial with $r = 2$ and $\beta = 1$.

For the sum of 1000 independent drivers:

Neg. Binomial with $r = (1000)(2) = 2000$ and $\beta = 1$, with mean $r\beta = 2000$, variance $r\beta(1 + \beta) = 4000$.

Prob[more than 2020 claims] \cong

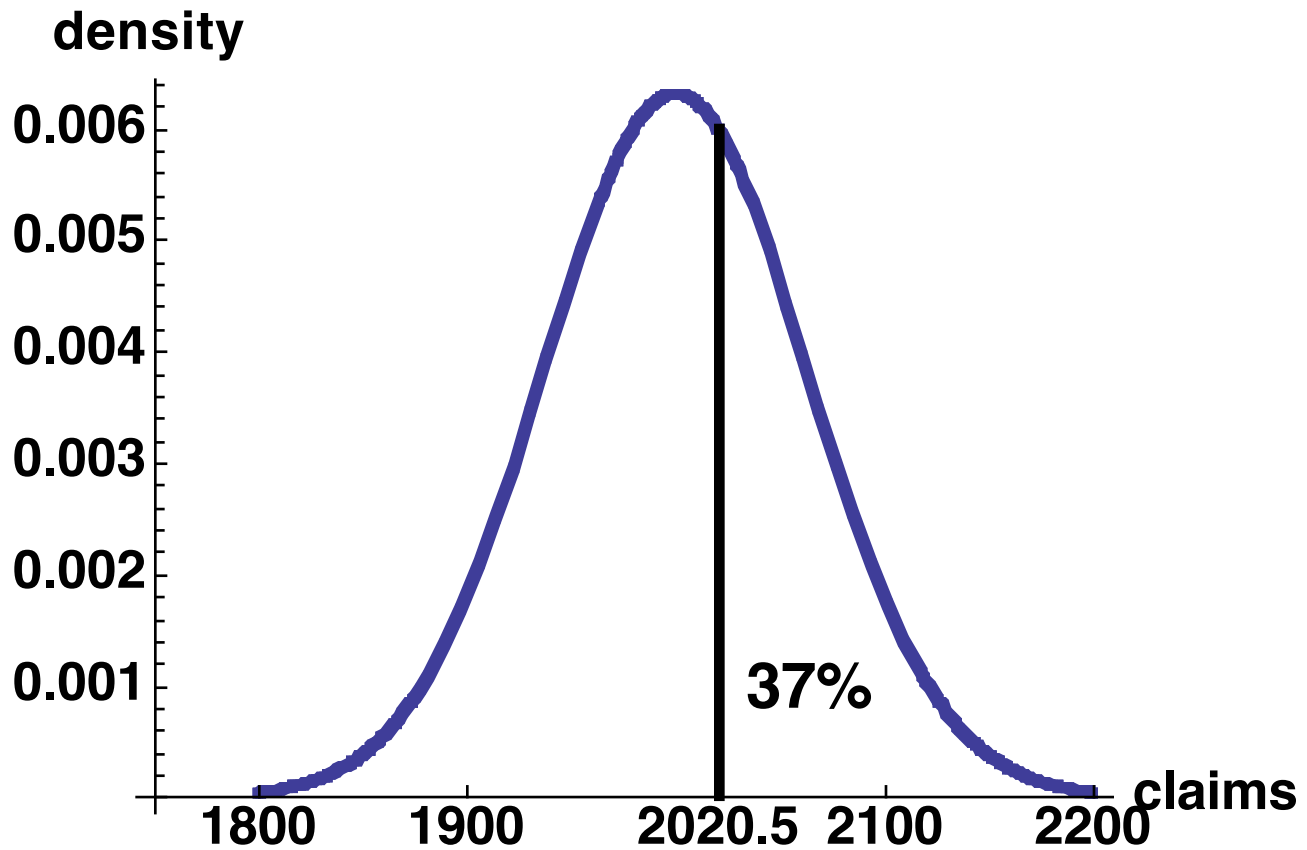
$$1 - \Phi\left[\frac{2020.5 - 2000}{\sqrt{4000}}\right] = 1 - \Phi[0.32] = \mathbf{37.45\%}.$$

Alternately, the mean of the sum of 1000 independent drivers is 1000 times the mean of single driver: $(1000)(2) = 2000$.

The variance of the sum of 1000 independent drivers is 1000 times the variance of single driver: $(1000)(2)(1)(1+1) = 4000$.

Proceed as before.

The Normal Distribution with mean 2000 and variance 4000, approximating the total number of claims from all 1000 drivers:



Tails of Frequency Distributions

Unlikely to be asked about on your exam.

Additional Questions

- For each individual driver, the number of accidents in a year follows a Poisson Distribution.
- For each individual driver, the mean of their Poisson Distribution λ is the same each year.
- For each individual driver, the number of accidents each year is independent of other years.
- The number of accidents for different drivers are independent.
- λ varies between drivers via a Gamma Distribution with mean 0.08 and variance 0.0032.
- Moe, Larry, and Curly are each drivers.

19.32. What is the probability that Moe has exactly one accident next year?

- A. 6.9% B. 7.1% C. 7.3% D. 7.5% E. 7.7%

19.32. B. For the Gamma, mean = $\alpha \theta = 0.08$,
and variance = $\alpha \theta^2 = 0.0032$.

Thus $\theta = 0.04$ and $\alpha = 2$.

This is a Gamma-Poisson, with mixed distribution
a Negative Binomial:

with $r = \alpha = 2$ and $\beta = \theta = 0.04$.

$$f(1) = \frac{r \beta}{(1 + \beta)^{r+1}} = \frac{(2)(0.04)}{(1 + 0.04)^3} = \mathbf{7.11\%}.$$

Comment: The fact that it is the next year rather
than some other year is irrelevant.

- For each individual driver, the number of accidents in a year follows a Poisson Distribution.
- For each individual driver, the mean of their Poisson Distribution λ is the same each year.
- For each individual driver, the number of accidents each year is independent of other years.
- The number of accidents for different drivers are independent.
- λ varies between drivers via a Gamma Distribution with mean 0.08 and variance 0.0032.
- Moe, Larry, and Curly are each drivers.

19.33. What is the probability that Larry has exactly 2 accidents over the next 3 years?

- A. 2.25% B. 2.50% C. 2.75% D. 3.00% E. 3.25%

19.33. C. For one year, each insured's mean is λ , and is distributed via a Gamma with:

$$\theta = 0.04 \text{ and } \alpha = 2.$$

Over three years, each insured's mean is 3λ , and is distributed via a Gamma with:

$$\theta = (3)(0.04) = 0.12, \text{ and } \alpha = 2.$$

This is a Gamma-Poisson, with mixed distribution a Negative Binomial:

with $r = \alpha = 2$ and $\beta = \theta = 0.12$.

$$\begin{aligned} f(2) &= \frac{r(r+1)}{2} \frac{\beta^2}{(1+\beta)^{r+2}} \\ &= \frac{(2)(3)}{2} \frac{0.12^2}{(1+0.12)^4} = \mathbf{2.75\%}. \end{aligned}$$

Comment: Assume a Gamma-Poisson model of insured drivers.

Lois has a low expected annual claim frequency, and Hi has a very high expected annual claim frequency.

Drivers such as Lois with a low λ in one year are assumed to have the same low λ every year.

Such good drivers have a small chance of having a large number of claims over several years.

Drivers such as Hi with a very high λ in one year are assumed to have the same high λ every year.

Such very bad drivers have a significant chance of having a large number of claims over several years.

Observe a Gamma-Poisson process for Y years, and each insured's Poisson parameter does not change over time.

$$\lambda \sim \text{Gamma}(\alpha, \theta).$$

Over Y years an insured is Poisson with mean $Y\lambda$.

$$Y\lambda \sim \text{Gamma}(\alpha, Y\theta), \text{ as per inflation.}$$

If one has a Poisson Distribution mixed by a Gamma Distribution with parameters α and θ , then over a period of length Y , the mixed distribution is Negative Binomial with $r = \alpha$ and $\beta = Y\theta$.

- For each individual driver, the number of accidents in a year follows a Poisson Distribution.
- For each individual driver, the mean of their Poisson Distribution λ is the same each year.
- For each individual driver, the number of accidents each year is independent of other years.
- The number of accidents for different drivers are independent.
- λ varies between drivers via a Gamma Distribution with mean 0.08 and variance 0.0032.
- Moe, Larry, and Curly are each drivers.

19.34. What is the probability that Moe, Larry, and Curly have a total of exactly 2 accidents during the next year?

- A. 2.25% B. 2.50% C. 2.75% D. 3.00% E. 3.25%

19.34. B. For one year, each insured's mean is λ , and is distributed via a Gamma with:

$$\theta = 0.04 \text{ and } \alpha = 2.$$

This is a Gamma-Poisson, with mixed distribution a Negative Binomial:

$$\text{with } r = \alpha = 2 \text{ and } \beta = \theta = 0.04.$$

We add up three individual independent drivers and we get a Negative Binomial with:

$$\text{with } r = \alpha = (3)(2) = 6, \text{ and } \beta = 0.04.$$

$$\begin{aligned} f(2) &= \frac{r(r+1)}{2} \frac{\beta^2}{(1+\beta)^{r+2}} \\ &= \frac{(6)(7)}{2} \frac{0.04^2}{(1+0.04)^8} = \mathbf{2.46\%}. \end{aligned}$$

Comment: The Negative Binomial Distributions here and in the previous solution have the same mean, however the densities are not the same.

- For each individual driver, the number of accidents in a year follows a Poisson Distribution.
- For each individual driver, the mean of their Poisson Distribution λ is the same each year.
- For each individual driver, the number of accidents each year is independent of other years.
- The number of accidents for different drivers are independent.
- λ varies between drivers via a Gamma Distribution with mean 0.08 and variance 0.0032.
- Moe, Larry, and Curly are each drivers.

19.35. What is the probability that Moe, Larry, and Curly have a total of exactly 3 accidents during the next four years?

- A. 5.2% B. 5.4% C. 5.6% D. 5.8% E. 6.0%

19.35. E. For one year, each insureds mean is λ , and is distributed via a Gamma with:

$$\theta = 0.04 \text{ and } \alpha = 2.$$

Over four years, each insureds mean is 4λ , and is distributed via a Gamma with:

$$\theta = (4)(0.04) = 0.16, \text{ and } \alpha = 2.$$

This is a Gamma-Poisson, with mixed distribution a Negative Binomial:

$$\text{with } r = \alpha = 2 \text{ and } \beta = \theta = 0.16.$$

We add up three individual independent drivers and we get a Negative Binomial with:

$$\text{with } r = \alpha = (3)(2) = 6, \text{ and } \beta = 0.16.$$

$$\begin{aligned} f(3) &= \frac{r (r + 1) (r + 2)}{6} \frac{\beta^3}{(1 + \beta)^{r + 3}} \\ &= \frac{(6) (7) (8)}{6} \frac{0.16^3}{(1 + 0.16)^9} = \mathbf{6.03\%}. \end{aligned}$$

V and X are each given by the result of rolling a six-sided die.

V and X are independent of each other.

$$Y = V + X.$$

$$Z = 2X.$$

Hint: The mean of X is 3.5

and the variance of X is $35/12$.

2.9. What is the standard deviation of Y?

A. less than 2.0

B. at least 2.0 but less than 2.3

C. at least 2.3 but less than 2.6

D. at least 2.9 but less than 3.2

E. at least 3.2

$$\mathbf{2.9. C.} \quad \text{Var}[Y] = \text{Var}[V + X] = \text{Var}[V] + \text{Var}[X] \\ = (35/12) + (35/12) = 35/6 = 5.83.$$

$$\text{Standard Deviation}[Y] = \sqrt{5.83} = \mathbf{2.41}.$$

V and X are each given by the result of rolling a six-sided die.

V and X are independent of each other.

$$Y = V + X.$$

$$Z = 2X.$$

Hint: The mean of X is 3.5

and the variance of X is $35/12$.

2.10. What is the standard deviation of Z?

A. less than 2.0

B. at least 2.0 but less than 2.3

C. at least 2.3 but less than 2.6

D. at least 2.9 but less than 3.2

E. at least 3.2

$$\begin{aligned} \mathbf{2.10. E.} \quad \text{Var}[Z] &= \text{Var}[2X] = 2^2 \text{Var}[X] \\ &= (4) (35/12) = 35/3 = 11.67. \end{aligned}$$

$$\text{Standard Deviation}[Z] = \sqrt{11.67} = \mathbf{3.42.}$$

14.28. X is a discrete random variable with a probability function which is a member of the $(a, b, 1)$ class of distributions.

p_k denotes the probability that $X = k$.

$p_1 = 0.1637$, $p_2 = 0.1754$, and $p_3 = 0.1503$.

Calculate p_5 .

(A) 7.5% (B) 7.7% (C) 7.9% (D) 8.1% (E) 8.3%

14.28. B.

Since we have a member of the $(a, b, 1)$ family:

$$p_2/p_1 = a + b/2.$$

$$\Rightarrow 2a + b = (2)(0.1754)/0.1637 = 2.1429.$$

$$p_3/p_2 = a + b/3.$$

$$\Rightarrow 3a + b = (3)(0.1503)/0.1754 = 2.5707.$$

$$\Rightarrow a = 0.4278. \Rightarrow b = 1.2873.$$

$$p_4 = (a + b/4) p_3$$

$$= (0.4278 + 1.2873/4) (0.1503) = 0.1127.$$

$$p_5 = (a + b/5) p_4$$

$$= (0.4278 + 1.2873/5) (0.1127) = \mathbf{0.0772}.$$

Comment: Based on a zero-modified Negative

Binomial, with $r = 4$, $\beta = 0.75$, and $p_0^M = 20\%$.

Important Formulas and Ideas

Here are what I believe are the most important formulas and ideas from this study guide to know for the exam.

Basic Concepts (Section 2)

The mean is the average or expected value of the random variable.

The mode is the point at which the density function reaches its maximum.

The median, the 50th percentile, is the first value at which the distribution function is ≥ 0.5 .

The 100pth percentile as the first value at which the distribution function $\geq p$.

Variance = second central moment = $E[(X - E[X])^2] = E[X^2] - E[X]^2$.

Standard Deviation = Square Root of Variance.

Binomial Distribution (Section 3)

$$f(x) = \binom{m}{x} q^x (1-q)^{m-x} = \frac{m!}{x! (m-x)!} q^x (1-q)^{m-x}, 0 \leq x \leq m.$$

Mean = mq

Variance = $mq(1-q)$

Probability Generating Function: $P(z) = \{1 + q(z-1)\}^m$

The Binomial Distribution for $m=1$ is a Bernoulli Distribution.

X is Binomial with parameters q and m_1 , and Y is Binomial with parameters q and m_2 ,

X and Y independent, then $X + Y$ is Binomial with parameters q and $m_1 + m_2$.

Poisson Distribution (Section 4)

$$f(x) = \frac{\lambda^x e^{-\lambda}}{x!}, x \geq 0$$

Mean = λ

Variance = λ

Probability Generating Function: $P(z) = e^{\lambda(z-1)}, \lambda > 0$.

A Poisson is characterized by a constant independent claim intensity and vice versa.

The sum of two independent variables each of which is Poisson with parameters λ_1 and λ_2 is also

Poisson, with parameter $\lambda_1 + \lambda_2$.

If frequency is given by a Poisson and severity is independent of frequency, then the number of claims above a certain amount (in constant dollars) is also a Poisson.

Geometric Distribution (Section 5)

$$f(x) = \frac{\beta^x}{(1+\beta)^{x+1}}$$

$$\text{Mean} = \beta$$

$$\text{Variance} = \beta(1+\beta)$$

$$\text{Probability Generating Function: } P(z) = \frac{1}{1 - \beta(z-1)}$$

For a Geometric Distribution, for $n > 0$, the chance of at least n claims is: $\left(\frac{\beta}{1+\beta}\right)^n$.

The Geometric shares with the Exponential distribution, the “memoryless property.” If one were to truncate and shift a Geometric Distribution, then one obtains the same Geometric Distribution.

Negative Binomial Distribution (Section 6)

$$f(x) = \frac{r(r+1)\dots(r+x-1)}{x!} \frac{\beta^x}{(1+\beta)^{x+r}} \quad \text{Mean} = r\beta \quad \text{Variance} = r\beta(1+\beta)$$

Negative Binomial for $r = 1$ is a Geometric Distribution.

For the Negative Binomial Distribution with parameters β and r , with r integer, can be thought of as the sum of r independent Geometric distributions with parameter β .

If X is Negative Binomial with parameters β and r_1 , and Y is Negative Binomial with parameters β and r_2 , X and Y independent, then $X + Y$ is Negative Binomial with parameters β and $r_1 + r_2$.

Normal Approximation (Section 7)

In general, let μ be the mean of the frequency distribution, while σ is the standard deviation of the frequency distribution, then the chance of observing at least i claims and not more than j claims is approximately:

$$\Phi\left[\frac{(j+0.5) - \mu}{\sigma}\right] - \Phi\left[\frac{(i-0.5) - \mu}{\sigma}\right].$$

Normal Distribution

$$F(x) = \Phi[(x-\mu)/\sigma]$$

$$f(x) = \phi[(x-\mu)/\sigma] / \sigma = \frac{\exp[-\frac{(x-\mu)^2}{2\sigma^2}]}{\sigma\sqrt{2\pi}}, -\infty < x < \infty. \quad \phi(x) = \frac{\exp[-x^2/2]}{\sqrt{2\pi}}, -\infty < x < \infty.$$

Mean = μ

Variance = σ^2

Skewness = 0 (distribution is symmetric)

Kurtosis = 3

Skewness (Section 8)

$$\begin{aligned} \text{Skewness} &= \text{third central moment} / \text{STDDEV}^3 = E[(X - E[X])^3] / \text{STDDEV}^3 \\ &= \{E[X^3] - 3\bar{X}E[X^2] + 2\bar{X}^3\} / \text{Variance}^{3/2}. \end{aligned}$$

A symmetric distribution has zero skewness.

Binomial Distribution with $q < 1/2 \Leftrightarrow$ positive skewness \Leftrightarrow skewed to the right.

Binomial Distribution $q = 1/2 \Leftrightarrow$ symmetric \Rightarrow zero skewness.

Binomial Distribution $q > 1/2 \Leftrightarrow$ negative skewness \Leftrightarrow skewed to the left.

Poisson and Negative Binomial have positive skewness.

Probability Generating Function (Section 9)

Probability Generating Function, p.g.f.:

$$P(z) = \text{Expected Value of } z^n = E[z^n] = \sum_{n=0}^{\infty} f(n) z^n.$$

The Probability Generating Function of the sum of independent frequencies is the product of the individual Probability Generating Functions.

The distribution determines the probability generating function and vice versa.

$$f(n) = \left(\frac{d^n P(z)}{dz^n} \right)_{z=0} / n! \quad f(0) = P(0). \quad P'(1) = \text{Mean}.$$

If a distribution is infinitely divisible, then if one takes the probability generating function to any positive power, one gets the probability generating function of another member of the same family of distributions. Examples of infinitely divisible distributions include: Poisson, Negative Binomial, Compound Poisson, Compound Negative Binomial, Normal, Gamma.

Factorial Moments (Section 10)

nth factorial moment $= \mu_{(n)} = E[X(X-1) \dots (X+1-n)]$.

$$\mu_{(n)} = \left(\frac{d^n P(z)}{dz^n} \right)_{z=1} \cdot \quad P'(1) = E[X]. \quad P''(1) = E[X(X-1)].$$

(a, b, 0) Class of Distributions (Section 11)

For each of these three frequency distributions: $f(x+1) / f(x) = a + \{b / (x+1)\}$, $x = 0, 1, \dots$ where a and b depend on the parameters of the distribution:

<u>Distribution</u>	<u>a</u>	<u>b</u>	<u>f(0)</u>
Binomial	$-q/(1-q)$	$(m+1)q/(1-q)$	$(1-q)^m$
Poisson	0	λ	$e^{-\lambda}$
Negative Binomial	$\beta/(1+\beta)$	$(r-1)\beta/(1+\beta)$	$1/(1+\beta)^r$

<u>Distribution</u>	<u>Mean</u>	<u>Variance</u>	<u>Variance Over Mean</u>	
Binomial	mq	$mq(1-q)$	$1-q < 1$	Variance < Mean
Poisson	λ	λ	1	Variance = Mean
Negative Binomial	$r\beta$	$r\beta(1+\beta)$	$1+\beta > 1$	Variance > Mean

<u>Distribution</u>	<u>Thinning by factor of t</u>	<u>Adding n independent, identical copies</u>
Binomial	$q \rightarrow tq$	$m \rightarrow nm$
Poisson	$\lambda \rightarrow t\lambda$	$\lambda \rightarrow n\lambda$
Negative Binomial	$\beta \rightarrow t\beta$	$r \rightarrow nr$

For X and Y independent:

<u>X</u>	<u>Y</u>	<u>X+Y</u>
Binomial(q, m ₁)	Binomial(q, m ₂)	Binomial(q, m ₁ + m ₂)
Poisson(λ ₁)	Poisson(λ ₂)	Poisson(λ ₁ + λ ₂)
Negative Binomial(β, r ₁)	Negative Bin.(β, r ₂)	Negative Bin.(β, r ₁ + r ₂)

Accident Profiles (Section 12)

For the Binomial, Poisson and Negative Binomial Distributions:

$(x+1) f(x+1) / f(x) = a(x + 1) + b$, where a and b depend on the parameters of the distribution.

a < 0 for the Binomial, a = 0 for the Poisson, and a > 0 for the Negative Binomial Distribution.

Thus if data is drawn from one of these three distributions, then we expect $(x+1) f(x+1) / f(x)$ for this data to be approximately linear with slope a; the sign of the slope, and thus the sign of a, distinguishes between these three distributions of the (a, b, 0) class.

Zero-Truncated Distributions (Section 13)

In general if f is a distribution on 0, 1, 2, 3, ..., then $p_k^T = \frac{f(k)}{1 - f(0)}$ is a distribution on 1, 2, 3,

<u>Distribution</u>	<u>Density of the Zero-Truncated Distribution</u>	
Binomial	$\frac{m! q^x (1-q)^{m-x}}{x! (m-x)!}$	$x = 1, 2, 3, \dots, m$
Poisson	$\frac{e^{-\lambda} \lambda^x / x!}{1 - e^{-\lambda}}$	$x = 1, 2, 3, \dots$
Negative Binomial	$\frac{r(r+1)\dots(r+x-1)}{x!} \frac{\beta^x}{(1+\beta)^{x+r}}$	$x = 1, 2, 3, \dots$

The moments of a zero-truncated distribution are given in terms of those of the corresponding untruncated

distribution, f, by: $E^{\text{Truncated}}[X^n] = \frac{E_f[X^n]}{1 - f(0)}$.

The Logarithmic Distribution has support on the positive integers: $f(x) = \frac{\left(\frac{\beta}{1+\beta}\right)^x}{x \ln(1+\beta)}$.

The **(a,b,1) class of frequency distributions** is a generalization of the (a,b,0) class.

As with the (a,b,0) class, the recursion formula applies: $\frac{\text{density at } x+1}{\text{density at } x} = a + \frac{b}{x+1}$.

However, it need only apply now for $x \geq 1$, rather than $x \geq 0$.

Members of the (a,b,1) family include: all the members of the (a,b,0) family, the zero-truncated versions of those distributions: Zero-Truncated Binomial, Zero-Truncated Poisson, Extended Truncated Negative Binomial, and the Logarithmic Distribution.

In addition the (a,b,1) class includes the zero-modified distributions corresponding to these.

Zero-Modified Distributions (Section 14)

If f is a distribution on $0, 1, 2, 3, \dots$, and $0 < p_0^M < 1$,

then probability at zero is p_0^M , $p_k^M = f(k) \frac{1 - p_0^M}{1 - f(0)}$, $k = 1, 2, 3, \dots$ is a distribution on $0, 1, 2, 3, \dots$

The moments of a zero-modified distribution are given in terms of those of f by:

$$E^{\text{Modified}}[X^n] = (1 - p_0^M) \frac{E_f[X^n]}{1 - f(0)} = (1 - p_0^M) E^{\text{Truncated}}[X^n].$$

Compound Frequency Distributions (Section 15)

A compound frequency distribution has a primary and secondary distribution, each of which is a frequency distribution. The primary distribution determines how many independent random draws from the secondary distribution we sum.

p.g.f. of compound distribution = p.g.f. of primary dist. [p.g.f. of secondary dist.]

$$P(z) = P_1[P_2(z)].$$

compound density at 0 = p.g.f. of the primary at the density at 0 of the secondary.

Moments of Compound Distributions (Section 16)

Mean of Compound Dist. = (Mean of Primary Dist.)(Mean of Sec. Dist.)

Variance of Compound Dist. = (Mean of Primary Dist.)(Var. of Sec. Dist.)

+ (Mean of Secondary Dist.)²(Variance of Primary Dist.)

In the case of a Poisson primary distribution with mean λ , the variance of the compound distribution could be rewritten as: $\lambda(2\text{nd moment of Second. Dist.})$.

The third central moment of a compound Poisson distribution = $\lambda(3\text{rd moment of Sec. Dist.})$.

Mixed Frequency Distributions (Section 17)

The density function of the mixed distribution, is the mixture of the density function for specific values of the parameter that is mixed.

The nth moment of a mixed distribution is the mixture of the nth moments.

First one mixes the moments, and then computes the variance of the mixture from its first and second moments.

The Probability Generating Function of the mixed distribution, is the mixture of the probability generating functions for specific values of the parameter.

For a mixture of Poissons, the variance is always greater than the mean.

Gamma Function (Section 18)

The (complete) **Gamma Function** is defined as:

$$\Gamma(\alpha) = \int_0^{\infty} t^{\alpha-1} e^{-t} dt = \theta^{-\alpha} \int_0^{\infty} t^{\alpha-1} e^{-t/\theta} dt, \text{ for } \alpha \geq 0, \theta \geq 0.$$

$$\Gamma(\alpha) = (\alpha-1)!$$

$$\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)$$

$$\int_0^{\infty} t^{\alpha-1} e^{-t/\theta} dt = \Gamma(\alpha)\theta^{\alpha}.$$

The **Incomplete Gamma Function** is defined as:

$$\Gamma(\alpha; x) = \int_0^x t^{\alpha-1} e^{-t} dt / \Gamma(\alpha).$$

Gamma-Poisson Frequency Process (Section 19)

If one mixes Poissons via a Gamma, then the mixed distribution is in the form of the Negative Binomial distribution with $r = \alpha$ and $\beta = \theta$.

If one mixes Poissons via a Gamma Distribution with parameters α and θ , then over a period of length Y , the mixed distribution is Negative Binomial with $r = \alpha$ and $\beta = Y\theta$.

For the Gamma-Poisson, the variance of the mixed Negative Binomial is equal to: mean of the Gamma + variance of the Gamma.

$\text{Var}[X] = E[\text{Var}[X \mid \lambda]] + \text{Var}[E[X \mid \lambda]]$. **Mixing increases the variance.**

Tails of Frequency Distributions (Section 20)

From lightest to heaviest tailed, the frequency distribution in the $(a,b,0)$ class are: Binomial, Poisson, Negative Binomial $r > 1$, Geometric, Negative Binomial $r < 1$.