Probability with Applications

Boon Han

May 14, 2017

Contents

Edit Log (Changes from Original PDF)

1. Sec 2.4(3) from $+P(EF)$ to $-P(EF)$
2. Sec 3 from $P(F) < 0$ to $P(F) > 0$
3. Sec 3.2 from "iare" to "are"
4. Sec 5.3 $F(a)$, added negative to exponent power
5. Sec 8 Markov's Inequality, from $\leq$ to $\geq$
6. Sec 5.4, Hazard Rate Function, changed t to t' and the bounds of the integral
7. Sec 10.2, Calculating Steady State probability vector, denominator of 1. removed unnecessary d_0.
8. Sec 10.1, added some information
9. Sec 5.2, added the case when normal random variable is a good approximation of a binomial
10. Added Conditional Variance, Section 7.4
11. Added Independence when dealing with conditional probabilities
12. Added Section 11, Entropy Equations

1 Combinatorial Analysis

1.1 The Generalized Principle of Counting

A way to calculate the number of outcomes of n_i experiments.

The Generalized Principle of Counting
We do r experiments. If the first will have n_1 possible outcomes and, for each of these n_1 outcomes there are n_2 outcomes of the second experiment, and for each possible outcome of the first two experiments there are n_3 possible outcomes... then there are a total of n_1 . n_2 ... n_r total possible outcomes of the r experiments.

1.2 Permutations

The number of ways to arrange n objects where order matters.

We can apply the generalized Principle of Counting to calculate the number of ways we can arrange n objects. If, of n objects, where n_1 are alike and n_2 are alike etc... then the number of ways to arrange them is

$$\frac{n!}{n_1!n_2!...n_r!}$$

Permutations of n objects taken r at a time
If we have n objects taken r at a time, then
1st object has n outcomes;
2nd object has *n-1* outcomes;
3rd object has *n-2* outcomes;
...
r^{th} object has *n-r+1* outcomes.
So the number of permutations of n objects taken r at a time is

$$P_{n,r} = \frac{n!}{(n-r)!}$$

1.3 Combinations

Permutations where order doesn't matter.

We can extend the concept of permuting n objects taken r at a time when the ordering within r does not matter. We remove repeated outcomes from $P_{n,r}$ to get the **Combination $\mathbf{C_{n,r}}$**:

$$C_{n,r} = \binom{n}{r} = \frac{n!}{r!(n-r)!}$$

A useful combinatorial identity is

$$\binom{n}{r} = \binom{n-1}{r-1} + \binom{n-1}{r} \qquad 1 \leq r \leq n$$

Which states that, if we take an arbitrary object o, then there are $\binom{n-1}{r-1}$ groups of size r that contain object o and $\binom{n-1}{r}$ groups of size r that do not contain object o.

1.4 The Binomial Theorem

The Binomial Theorem

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

1.5 Multinomial Coefficients

Multinomial Coefficients

If $n_1 + n_2 + .. + n_r = n$ then we define:

$$\binom{n}{n_1, n_2, ..., n_r} = \frac{n!}{n_1!n_2!..n_r!}$$

which represents the number of possible divisions of n **distinct** objects into distinct groups of sizes n_1, n_2,...,n_r.

The Multinomial Theorem

$$(x_1 + x_2 + ... + x_r)^n = \sum_{(n_1,...n_r):n_1+n_2+...+n_r=n} \binom{n}{n_1, n_2, ..., n_r} x_1^{n_1} x_2^{n_2} ... x_r^{n_r}$$

The subtext of the summation means all values of n_i have to sum up to n.

1.6 Number of Integer Solutions of Equations (Bars and Stars Method)

Here we find a way to find the number of ways to satisfy the equation $x_1 + x_2 + ... + x_r = n$. This is the same problem as trying to divide n indistinguishable objects into r nonempty groups. Since we choose $r-1$ of the $n-1$ possible spaces between objects there are

$$\binom{n-1}{r-1}$$

possible selections which are **distinctly positive**. (IE $x_i > 0$) If we want **distinct non-negative** values, then we find the solution to $y_1, y_2, ...y_r = n + r$ where $y_i = x_i + 1, i = 1, .., r$ and this gives us:

$$\binom{n+r-1}{r-1}$$

If not all of n needs to be "used", then we can add a new variable on the LHS which will denote the amount of n "left out".

1.7 Additional Information

Order Matters Repetition Allowed	n^r
Order Matters Repetition Not Allowed	$\frac{n!}{(n-r)!}$
Order Doesn't Matter Repetition Allowed	$\frac{n!}{r!(n-r)!}$
Order Doesn't Matter Repetition Not Allowed	$\frac{(n+r-1)!}{(n-1)!}$

2 Axioms of Probability

Here we talk about the concept of probability of an event and show how probabilities can be computed.

2.1 Sample Space

We call the set of all possible outcomes of an experiment as the *sample space* of an experiment. For example, for the experiment of the gender of a unborn child, the outcome sample space S is

$$S = \{g, b\}$$

Any subset of this sample space is known as an *event*. So if E is the event that a girl is born, we can describe it as

$$E = \{g\}$$

More complex events can be described using the $\cup$ union or $\cap$ intersection symbols, where $E \cup F$ describes the event where either E and F occurs, while $E \cap F$ describes the event where both E and F occurs. For more than two events we have the equivalent $\bigcup_{n=1}^{\infty} E_n$ and $\bigcap_{n=1}^{\infty} E_n$ respectively.

2.2 De Morgan's Laws

De Morgan's Laws

$$\left(\bigcup_{i=1}^{\infty} E_i \right)^c = \bigcap_{i=1}^{\infty} E_i^c$$

$$\left(\bigcap_{i=1}^{\infty} E_i \right)^c = \bigcup_{i=1}^{\infty} E_i^c$$

2.3 Probability of an Event

$$P(E) = \lim_{n \to \infty} \frac{n(E)}{n}$$

2.4 Simple Propositions

Here some propositions are listed.

$$P(E^c) = 1 - P(E) \tag{1}$$

$$\text{If } E \subset F \text{ then } P(E) \leq P(F) \tag{2}$$

$$P(E \cup F) = P(E) + P(F) - P(EF) \tag{3}$$

$$P(E_1 \cup E_2 ... \cup E_n) = \sum_{i=1}^{n} P(E_i) - \sum_{i_1 < i_2} P(E_{i_1} E_{i_2}) + ... \tag{4}$$

$$+ (-1)^{r+1} \sum_{i_1 < i_2 ... < i_r} P(E_{i_1} E_{i_2} ... E_{i_r})$$

$$+ ... + (-1)^{n+1} P(E_1 E_2 ... E_n)$$

An example of (4) is $P(E \cup F \cup G) = P(E) + P(F) + P(G) - P(EF) - P(EG) - P(FG) + P(EFG)$. Note that if we **assume equal probabilities** then we can simplify to something like $3.P(E) - 3.P(EF) + P(EFG)$. Note that due to the flipping nature of the addition / subtraction signs, the first term gives an upper bound on $P(E)$, the second gives a lower bound, the third gives upper, etc etc.

2.5 Equally Likely Outcomes

If we assume that all outcomes in a sample space S are equally likely then we can conclude that

$$P(E) = \frac{n(E)}{n(S)}$$

3 Conditional Probability and Independence

Conditional Probability
If $P(F) > 0$, then

$$P(E|F) = \frac{P(EF)}{P(F)}$$

The Multiplication Rule

$$P(E_1E_2...E_n) = P(E_1)P(E_2|E_1)P(E_3|E_1E_2)...P(E_n|E_1...E_{n-1})$$

3.1 Baye's Formula

Sometimes we cannot calculate the Probability of an event E directly. We instead calculate the conditional probability of E occuring based on whether or not a second event F has occured. Baye's Formula then allows us to derive $P(E)$ because:

Baye's Formula

$$P(E) = P(E|F)P(F) + P(E|F^c)(1 - P(F))$$

If we we want to use this information to support a particular hypothesis, IE say that **This information E makes it more likely that an event has occured** then necessarily some conditions must hold. We know that

$$P(H|E) = \frac{P(E|H)P(H)}{P(H)P(E|H) + P(H^c)P(E|H^c)}$$
$$P(E|H) \uparrow \iff P(E|H) \geq P(H)P(E|H) + P(H^c)P(E|H^c)$$
$$\boldsymbol{P(E|H) \geq P(E|H^c)}$$

In other words, an event H is more likely after factoring some information E if E is more likely given H than H^c. So if it is more likely $[P(E|H) \geq P(E|H^c)]$ to have cancer (H) if you are a smoker (E) than not (E^c), then knowing that someone smokes increases the the probability that this person will have cancer. $[P(H|E) \uparrow]$
In fact the change in probability of a hypothesis when a new evidence is introduced can be express compactly as the *odds* of a hypothesis.

Odds of an Event A

$$\alpha = \frac{P(A)}{P(A^c)} = \frac{P(A)}{1 - P(A)}$$

Tells us how more likely it is that an event A occurs than it does not. If the odds are equal to α, then we say that odds are "α to 1" in favour of the hypothesis.
The new odds after evidence E is known, given a hypothesis H, is then a function of the old odds:

$$\frac{P(H|E)}{P(H^c|E)} = \frac{P(H)}{P(H^c)}\frac{P(E|H)}{P(E|H^c)}$$

3.2 Indepedent Events

We know from before that $P(E|F) = \frac{P(EF)}{P(F)}$. Generally we cannot simplify further. However, in the case that E and F do not depend on each other, we know that $P(E|F) = P(E)$. This can only happen if E and F are **independent**.

Independence

If two events are independent, then $P(EF) = P(E)P(F)$

It is also true that if E and F are independent, then so are E and F^c

If we have n terms , these n terms are independent only if every subset of these terms are independent. For example,

Independence of Three Events

Three events E, F, G are independent only if

$$P(EFG) = P(E)P(F)P(G)$$
$$P(EF) = P(E)P(F)$$
$$P(FG) = P(F)P(G)$$
$$P(EG) = P(E)P(G)$$

Note that there are 2^n subsets, so if we exclude subsets of size 1 (trivially, $P(E) = P(E)$) and size 0 (null set), we have $2^n - n - 1$ subsets to consider.

4 Random Variables

A random variable X is a variable whose possible values describe the possible numeric outcomes of a random phenomenon. So, for a dice roll, $P(X = 1) = \frac{1}{6}$.

Cumulative Distribution Function

$$F(x) = P(X \leq x) \qquad -\infty < x < \infty$$

4.1 Discrete Random Variables

Probability Mass Function p

$$p(a) = P\{X = a\}$$

$p(a)$ is a *positive number* (not function) for a countable number of values of a, and is 0 otherwise. Note that because we are dealing with probabilities,

$$\sum_{i=1}^{\infty} p(x_i) = 1$$

Cumulative Distribution Function F

$$F(a) = \sum_{\text{all } x \leq a} p(x)$$

Note that $F(a)$ is a step function where the value is constant at intervals and then takes a jump at certain points.

Expected Value

$$E[X] = \sum_{x:p(x)>0} xp(x)$$

A weighted average of the possible values X can take, weighted by its probability.

4.2 Expectation of Function of Discrete Random Variable

Expectation of a Function of a Random Variable

$$E[g(X)] = \sum_i g(x_i)p(x_i)$$

Note that x_i is replaced with $g(x_i)$.
Consequently, we also can say that:

$$E[aX + b] = aE[X] + b$$

$$E[X^n] = \sum_{x:p(x)>0} x^n p(x)$$

4.3 Variance

A measure of the deviation of a random variable from its mean.

Variance

$$Var(X) = E[(X - \mu)^2]$$

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

$$Var(aX + b) = a^2 Var(X)$$

4.4 Bernoulli and Binomial Random Variables

Bernoulli Random Variable (n, p)
An single experiment whose outcome can be classified as either a *success* or *failure*.
The probability mass function $p(a)$ is given by

$$p(0) = P\{X = 0\} = 1 - p$$

$$p(1) = P\{X = 1\} = p$$

Binomial Random Variable ($n > 0$, $0 < p < 1$)
n independent bernoulli random variables with probability of success p are carried out. X represents **number of successes** in n trials.
The probability mass function $p(i)$ is given by:

$$p(i) = \binom{n}{i} p^i (1-p)^{n-i} \qquad i = 0, 1, ..., n$$

Properties of Binomial Random Variables

$$E[X] = np$$

$$Var(X) = np(1-p)$$

If X is a binomial Random Variable with parameters n and p, where $0 < p < 1$ then as k goes from 0 to n, $P\{X = k\}$ first increases monotonically and then decreases monotonically, reaching its largest value when k is the largest integer $\leq (n+1)p$.

Calculating the Binomial Distribution Function

To calculate the cumulative distribution function of a Binomial RV(n, p) $P\{X \leq i\}$, we use the known relationship:

$$P\{X = k+1\} = \frac{p}{1-p}\frac{n-k}{k+1}P\{X = k\}$$

We can start with $P\{X = 0\}$ and then derive the $k+1$ term recursively.

4.5 Poisson Random Variable

A good estimation to the Binomial Random Variable if n is large and p is small such that np is of moderate size. In this case, $\lambda = np$.

Poisson Random Variable($\lambda > 0$)

$$p(i) = P\{X = i\} = e^{-\lambda}\frac{\lambda^i}{i!} \qquad i = 0, 1, 2...$$

λ is also the Expected number of successes.

The following are some Random Variables that fit a Poisson RV. The common theme is that all of them have a large n but a small p.

1. The number of misprints on a book
2. Number of people in a community that live to 100
3. Number of wrong telephone numbers that are dialed a day
4. Number of packages of dog biscuits sold in a particular store every day
5. etc...

Computing the Poisson Random Variable

$$\frac{P\{X = i+1\}}{P\{X = i\}} = \frac{e^{-\lambda}\lambda^{i+1}/(i+1)!}{e^{-\lambda}\lambda^{i}/i!} = \frac{\lambda}{i+1}$$

and we can start with $P\{X = 0\} = e^{-\lambda}$ and compute successive $P\{X = a\}$ since

$$P\{X = i+1\} = \frac{\lambda}{i+1}P\{X = i\}$$

Properties of Poisson Random Variables

$$E(X) = \lambda$$

$$Var(X) = \lambda$$

4.6 Geometric Random Variable

A binomial random variable with consecutive failures and terminates on its first success. X is the number of trials required to get this success.

Geometric Random Variable $(0 < p < 1)$

$$P\{X = n\} = (1-p)^{n-1}p \qquad n = 1, 2, ...$$

Note that in the Binomial RV, n is fixed while i is the variable denoting number of successes. Here, n is the variable.

Properties of Geometric Random Variables

$$E(X) = \frac{1}{p}$$

$$Var(X) = \frac{1-p}{p^2}$$

4.7 Negative Binomial Random Variable

A Binomial Random Variable that terminates when r successes are accumulated. In essence, we have a Binomial Random Variable modelling the first $r-1$ successes, followed by a terminating success with probability p.

Negative Binomial Random Variable ($r \in \mathbb{R}^+$, p)

$$P\{X = n\} = \binom{n-1}{r-1} p^r (1-p)^{n-r} \qquad n = r, r+1...$$

Note that the **geometric random variable** is just a negative binomial with $r = 1$.

Propeties of a Negative Binomial

$$E(X) = \frac{r}{p}$$

$$Var(X) = \frac{r(1-p)}{p^2}$$

4.8 Hypergeometric Random Variable

A binomial RV requires independence between experiments. If p changes based on the ratio of two binary elements, such as when we have a urn of N balls, m white and $N-m$ black, the "successes" is described as follows:

Hypergeometric Random Variable (n, N, m)

$$P\{X = i\} = \frac{\binom{m}{i}\binom{N-m}{n-i}}{\binom{N}{n}} \qquad i = 0, 1, ..., n$$

Very simply, we use the generalized counting principle to count i "successes", $n-i$ failures conditioned on the fact that we drew n samples.

Properties of Hypergeometric Random Variables

$$E[X] = \frac{nm}{N}$$

$$Var(x) = np(1-p)(1 - \frac{n-1}{N-1})$$

$$Var(x) \approx np(1-p) \qquad \text{when N is large w.r.t n}$$

Intuitively, a Hypergeometric RV approximates a Binomial RV for large N w.r.t n because if the sample drawn is small, the probabilities *remain almost constant*

4.9 Expected Value of Sums of Discrete Random Variables

Linearity of Expectations

$$E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i]$$

4.10 Properties of the Discrete Cumulative Distribution Function

The Cumulative Distribution Function denotes the probability that a random variable takes on a value *less than or equal to b*.

1. F is non-decreasing function
2. $\lim_{b\to\infty} F(b) = 1$
3. $\lim_{b\to-\infty} F(b) = 0$
4. F is right continuous.
5. $P\{a < X \leq b\} = F(b) - F(a) \qquad \forall a, b.a < b$

5 Continuous Random Variables

As opposed to *discrete* random variables whose set of possible values is either finite or countably infinite, there exist random variables whose set of possible values is **uncountable**. This gives rise to *continuous random variables.* We are concerned with the probability that the RV is within a set of values B.

Probability Density Function f

$$f(x) = P\{X = a\}$$

where f is some continuous function.

Cumulative Distribution Function F

$$P\{X \in B\} = \int_B f(x)dx$$

$$P\{a \leq X \leq b\} = \int_a^b f(x)dx$$

$$P\{X < a\} = P\{X \leq a\} = F(a) = \int_{-\infty}^{a} f(x)dx$$

Note that $P\{X < a\} = P\{X \leq a\}$ because $\int_a^a f(x) = 0$.

Relationship between F and f

$$\frac{d}{da}F(a) = f(a)$$

The derivative of the Cumulative Distribution Function is the Probability Density.

Properties of Continuous Random Variables

$$E[X] = \int_{-\infty}^{\infty} xf(x)dx$$

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)dx$$

$$E[aX + b] = aE[X] + b$$

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

$$Var(aX + b) = a^2Var(X)$$

5.1 Uniform Random Variables

A continuous random variable is uniform if its probability within an interval is constant, and zero otherwise.

Uniform Random Variable (α, β)

A random variable is uniformly distributed over an interval (α, β) if

$$f(x) = \begin{cases} \frac{1}{\beta-\alpha} & \alpha < x < \beta \\ 0 & \text{otherwise} \end{cases}$$

$$F(a) = \begin{cases} 0 & a \leq \alpha \\ \frac{a-\alpha}{\beta-\alpha} & \alpha < x < \beta \\ 1 & \alpha \geq \beta \end{cases}$$

Properties of Uniform Random Variables

$$E[X] = \frac{\beta + \alpha}{2}$$

$$Var(X) = \frac{(\beta - \alpha)^2}{12}$$

5.2 Normal Random Variables

A random variable whose probability density follows a bell-curve shape. A good approximation of binomial when $N.p.(1-p) \geq 10$

Normal Random Variable (μ, σ^2)

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

Standardized Normal Random Variable

A normally distributed random variable X can be standardized: $Z = \frac{X-\mu}{\sigma}$ which is normally distributed with parameters $\mu = 0, \sigma^2 = 1$ such that

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2} dy$$

such that we can represent the Cumulative Distribution Function $F_X(a)$ as:

$$F_X(a) = P\{X \leq a\} = \Phi\left(\frac{a-\mu}{\sigma}\right)$$

Properties of a Normal Random Variable

$$E[X] = \mu$$

$$Var(X) = \sigma^2$$

$$E[aX + b] = a\mu + b$$

$$Var(aX + b) = \alpha^2\sigma^2$$

The following theorem is to estimate a binomial random variable using a normal approximation. This works when n is large.

The DeMoivre-Laplace Limit Theorem
If S_n denotes the number of successes in a binomial experiment then the "standardized" S_n is approximated by a standard normal random variable:

$$P\left\{a \leq \frac{S_n - np}{\sqrt{np(1-p)}} \leq b\right\} \to \Phi(b) - \Phi(a)$$

as n $\to \infty$.
Since we are applying a continuous RV to estimate a discrete RV, we apply *continuity correction.*

$$If P\{X = 20\} \to_{c.corr} P\{19.5 \leq X < 20.5\}$$

Then we **normalize** the S_n terms (19.5 and 20.5) and use the standardized normal RV Φ.

$$\Phi(b) - \Phi(a) = \Phi(\frac{19.5 - 20}{\sqrt{10}}) - \Phi(\frac{20.5 - 20}{\sqrt{10}})$$

5.3 Exponential Random Variables

An Exponential Random Variable is used to model the amount of time until a specific event occurs.

Exponential Random Variable (λ) The Exponential RV Probability Density is given as:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

$$F(a) = 1 - e^{-\lambda a} \qquad a \geq 0$$

Properties of Exponential Random Variables

$$E[X] = \frac{1}{\lambda}$$

$$Var(X) = \frac{1}{\lambda^2}$$

Memoryless: $P\{X > s + t | X > t\} = P\{X > s\} \quad \text{for all } s, t \geq 0$

Equivalently: $P\{X > s + t\} = P\{X > s\}P\{X > t\}$

5.4 Hazard Rate Functions

The Hazard Rate Function is the conditional probability that a t-unit-year-old item will fail at time t' given that it already survived up to t.
For example, the death rate of a person who smokes is, at each age, twice that of a non-smoker.

Survival Function

$$\text{Survival Function } = 1 - F(t)$$

Hazard Rate Functions
A ratio of the probability density function to survival function.

$$\lambda(t) = \frac{f(t)}{\bar{F}(t)}$$

The parameter λ of an exponential distribution is often referred to as the **rate** of the distribution.
The failure rate function for the exponential distribution uniquely determines the distribution F.
In general,

$$F(t) = 1 - exp\bigg\{ - \int_0^t \lambda(t)dt \bigg\}$$

For example, if a random variable has a hazard rate function

$$\lambda(t) = a + bt$$

then the cumulative distribution function $F(t)$ is $1 - e^{-(at+bt^2/2)}$ and differentiation will yield the density

$$f(t) = (a + bt)e^{-(at+bt^2/2)} \qquad t \geq 0$$

6 Jointly Distributed Random Variables

We now extend probability distributions to two random variables.

Jointly Distributed Random Variables

$$F(a,b) = P\{X \leq a, Y \leq b\} \qquad -\infty < a, b < \infty$$

We can obtain marginal distribution of X from the joint distribution F easily:

$$F_X(a) = F(a, \infty)$$

And we can answer all joint probability statements about X and Y because:

$$P\{a_i < X \leq a_2, b_1 < Y \leq b_2\} = F(a_2,b_2) + F(a_1,b_1) - F(a_1,b_2) - F(a_2,b_1)$$

Discrete Jointly Distributed Random Variables The discrete case is modelled with the *joint probability mass function* by

$$p(x,y) = P\{X = x, Y = y\}$$

and marginals (example of X, with trivial equivalent for Y)

$$p_X(x) = \sum_{y:p(x,y)>0} p(x,y)$$

Note that we can represent the joint probability mass function best in table form.

Continuous Jointly Distributed Random Variables The continuous case is modelled by (note there are two variables)

$$P\{(X,Y) \in C\} = \iint_{(x,y)\in C} f(x,y)\ dxdy$$

$$P\{X \in A, Y \in B\} = \int_B \int_A f(x,y)\ dxdy$$

$$F(a,b) = \int_{-\infty}^{b} \int_{-\infty}^{a} f(x,y)\ dxdy$$

$$f(a,b) = \frac{\partial^2}{\partial a\ \partial b} F(a,b)$$

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\ dy \qquad \text{with trivial equivalent for y}$$

6.1 Independent Random Variables

Here we discuss the situation where knowing the value of one Random Variable does not affect the other.

Independence of Jointly Distributed Random Variables Two Random Variables are *independent* if

$$P\{X \in A, Y \in B\} = P\{X \in A\}P\{Y \in B\}$$

Consequently, we can say

$$P\{X \leq a, Y \leq b\} = P\{X \leq a\}P\{Y \leq b\}$$

Hence, in terms of joint distribution, X and Y are independent if

$$F(a, b) = F_X(a)F_Y(b) \qquad \text{for all } a, b$$

For the discrete case we have

$$p(x, y) = p_X(x)p_Y(y) \qquad \text{for all } x, y$$

and for the continuous case we have

$$f(x, y) = f_X(x)f_Y(y) \qquad \text{for all } x, y$$

The continuous (discrete) random variables X and Y are independent if and only if their joint probability density(mass) functions can be expressed as follows:

$$f_{X,Y}(x, y) = h(x)g(y) \qquad -\infty < x < \infty, -\infty < y < \infty$$

Note that independent is symmetric. Consequently, if its difficult to determine independece in our direction, it is valid to go the other way.

Conditional Distribution: Discrete Case
Conditional Probabillity Mass Function of X given $Y = y$

$$p_{X|Y}(x|y) = \frac{p(x, y)}{p_Y(y)}$$

Conditional Probability Distribution Function of X given $Y = y$

$$F_{X|Y}(x|y) = \sum_{a \leq x} p_{X|Y}(a|y)$$

If X and Y are independent then then

$$p_{X|Y}(x|y) = P\{X = x\}$$

6.2 Conditional Distribution: Continuous Case

Conditional Density Function

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)}$$

Conditional Probabilities of Events

$$P\{X \in A|Y = y\} = \int_A f_{X|Y}(x|y)\ dx$$

Conditional Cumulative Distribution Function of X (with trivial equivalent for Y)

$$F_{X|Y}(a|y) = \int_{-\infty}^{a} f_{X|Y}(x|y)\ dx$$

A conditional probability such as $P\{X > 1|Y = y\}$ is a function with variables x and y.

6.3 Multinomial Distribution

A multinomial distribution arises when n independent and identical experiments are performed. We have n random variables, and we denote number of experiments with outcome i with the random variable X_i.

The Multinomial Distribution

$$P\{X_1 = n_1, X_2 = n_2 ... X_r = n_r\} = \frac{n!}{n_1!n_2!...n_r!} p_1^{n_1} p_2^{n_2} ... p_r^{n_r}$$

7 Properties of Expectation

7.1 Expectation of Sum of Random Variables

Discrete Case

$$E[g(X,Y)] = \sum_y \sum_x g(x,y)p(x,y)$$

Continuous Case

$$E[g(X,Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y)f(x,y)\ dxdy$$

7.2 Covariance, Variance of Sums, and Correlations

Expectation of Product of Independent Variables

If X and Y are independent, then

$$E[g(X)h(Y)] = E[h(Y)]E[g(X)]$$

Covariance between X and Y

The Covariance between X and Y gives us information about the relationships about the two RV.

$$\begin{aligned} Cov(X,Y) &= E[(X - E[X])(Y - E[Y])] \\ &= E[XY] - E[X]E[Y] \end{aligned}$$

Note that if X and Y are independent, then $Cov(X,Y) = 0$. However, the converse is not true.

There are some things to note:

1. $Cov(X,Y) = Cov(Y,X)$
2. $Cov(X,X) = Var(X)$
3. $Cov(aX,Y) = aCov(X,Y)$
4. $Cov\Bigg(\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j \Bigg) = \sum_{i=1}^{n} \sum_{j=1}^{m} Cov(X_i, Y_j)$

Variance of multiple Random Variables

In general,

$$Var\left(\sum_{i=1}^{n} X_i \right) = \sum_{i=1}^{n} Var(X_i) + 2 \sum \sum_{i<j} Cov(X_i, X_j)$$

If $X_1, ..., X_n$ are pairwise independent, meaning each event is independent of *every other possible combination* of paired events, then

$$Var\left(\sum_{i=1}^{n} X_i \right) = \sum_{i=1}^{n} Var(X_i)$$

Note that mutual independence is pairwise independence with the additional requirement that $P(A \cap B \cap ...) = P(A) \times P(B) \times ...$

Correlation

A measure of linearity between two Random Variables. A value near +1 or -1 indicates high degree of linearity, while a value near 0 indicates otherwise.

As long as $Var(X)Var(Y)$ is positive, the correlation between X and Y is defined as

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$$

and $-1 \leq \rho(X,Y) \leq 1$

7.3 Conditional Expectation

We extend Expectations to the case where we involve conditional probabilities.

Discrete Case

$$E[X|Y=y] = \sum_x x p_{X|Y}(x|y)$$

Continuous Case

$$E[X|Y=y] = \int_{-\infty}^{\infty} x f_{X|Y}(x|y)\ dx$$

Computing Expectations by Conditioning

If we denote $E[X|Y]$ to be the function of the random variable Y whose value at $Y=y$ is $E[X|Y=y$, then

$$E[X] = E[E[X|Y]]$$

In the discrete case,

$$E[X] = \sum_y E[X|Y=y]P\{Y=y\}]$$

$$E[X] = \int_{-\infty}^{\infty} E[X|Y=y]f_Y(y)\ dy$$

This allows us to easily compute expectations by first conditioning on some appropriate variable.

7.4 Conditional Variance

Conditional Variance

$$Var(X|Y) = E[(X - E[X|Y])^2|Y]$$

$$Var(X|Y) = E[X^2|Y] - (E[X|Y])^2$$
$$Var(X) = E[Var(X|Y)] + Var(E[X|Y])$$

7.5 Independence in Conditional Probabilities

Independence in Conditional Probability

$$E[X|Y] = E[X]$$

$$Var(X|Y) = Var(X)$$

7.6 Computing Probabilities by Conditioning

We extend the computation of probabilities to conditional statements.

Discrete Case

$$P(E) = \sum_y P(E|Y = y)P(Y = y)$$

Continuous Case

$$P(E) = \int_{-\infty}^{\infty} P(E|Y = y)f_Y(y)\ dy$$

7.7 Moment Generating Functions

Moments of a random variable X (Expectations of the powers of X) can be calculated using Moment Generating Functions:

$$M(t) = E[e^{tX}]$$

Discrete Case, X has mass function $p(x)$

$$M(t) = \sum_x e^{tx}p(x)$$

Continuous Case, X has density function $f(x)$

$$M(t) = \int_{-\infty}^{\infty} e^{tx}f(x)\ dx$$

We do this by differentiating n times and evaluating at $t = 0$ to get the expectation of the n^{th} moment. For Example,

$$M'(0) = E[X]$$

	Probability mass function, $p(x)$	Moment generating function, $M(t)$	Mean	Variance
Binomial with parameters n, p; $0 \le p \le 1$	$\binom{n}{x} p^x (1-p)^{n-x}$ $x = 0, 1, \ldots, n$	$(pe^t + 1 - p)^n$	np	$np(1-p)$
Poisson with parameter $\lambda > 0$	$e^{-\lambda} \dfrac{\lambda^x}{x!}$ $x = 0, 1, 2, \ldots$	$\exp\{\lambda(e^t - 1)\}$	λ	λ
Geometric with parameter $0 \le p \le 1$	$p(1-p)^{x-1}$ $x = 1, 2, \ldots$	$\dfrac{pe^t}{1-(1-p)e^t}$	$\dfrac{1}{p}$	$\dfrac{1-p}{p^2}$
Negative binomial with parameters r, p; $0 \le p \le 1$	$\binom{n-1}{r-1} p^r (1-p)^{n-r}$ $n = r, r+1, \ldots$	$\left[\dfrac{pe^t}{1-(1-p)e^t}\right]^r$	$\dfrac{r}{p}$	$\dfrac{r(1-p)}{p^2}$

	Probability mass function, $f(x)$	Moment generating function, $M(t)$	Mean	Variance
Uniform over (a, b)	$f(x) = \begin{cases} \dfrac{1}{b-a} & a < x < b \\ 0 & \text{otherwise} \end{cases}$	$\dfrac{e^{tb} - e^{ta}}{t(b-a)}$	$\dfrac{a+b}{2}$	$\dfrac{(b-a)^2}{12}$
Exponential with parameter $\lambda > 0$	$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \ge 0 \\ 0 & x < 0 \end{cases}$	$\dfrac{\lambda}{\lambda - t}$	$\dfrac{1}{\lambda}$	$\dfrac{1}{\lambda^2}$
Gamma with parameters $(s, \lambda), \lambda > 0$	$f(x) = \begin{cases} \dfrac{\lambda e^{-\lambda x}(\lambda x)^{s-1}}{\Gamma(s)} & x \ge 0 \\ 0 & x < 0 \end{cases}$	$\left(\dfrac{\lambda}{\lambda - t}\right)^s$	$\dfrac{s}{\lambda}$	$\dfrac{s}{\lambda^2}$
Normal with parameters (μ, σ^2)	$f(x) = \dfrac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2} \quad -\infty < x < \infty$	$\exp\left\{\mu t + \dfrac{\sigma^2 t^2}{2}\right\}$	μ	σ^2

8 Limit Theorems

Here we study some important conclusions of probability theory. Markov's and Chebyshev's inequalities allow us to derive bounds on probabilities when only the mean, or both mean and variance, of the probability distribution are known.

8.1 Markov's Inequality

If X is a random variable that takes only non-negative values then for any $a > 0$,

$$P\{X \geq a\} \leq \frac{E[X]}{a}$$

8.2 Chebyshev's Inequality

If X is a random variable with finite mean μ and variance σ^2 then for any $k \geq 0$,

$$P\{|X - \mu| \geq k\} \leq \frac{\sigma^2}{k^2}$$

$$P\{|X - \mu| \geq k\sigma\} \leq \frac{1}{k^2}$$

8.3 Chernoff's Bounds

If X is a random variable with $M_x(t) = E[e^{tX}]$ then

$$P(X \geq a) \leq e^{-ta} M_x(t) \qquad \text{for all } t \geq 0$$

$$P(X \leq a) \leq e^{-ta} M_x(t) \qquad \text{for all } t < 0$$

8.4 Weak Law of Large Numbers

The weak law of large numbers (also called Khintchine's law) states that the sample average converges in probability towards the expected value.
Let $X_1, X_2, ..., X_n$ be a sequence of independent and indentically distributed random variables, each with finite mean $E[X] = \mu$. Then for any $\epsilon > 0$,

$$P\left\{\left|\frac{X_1 + ... + X_n}{n} - \mu\right| \geq \epsilon\right\} \to 0 \qquad \text{as } n \to \infty$$

8.5 Strong Law of Large Numbers

Let $X_1, X_2, ..., X_n$ be a sequence of independent and indentically distributed random variables, each with finite mean $E[X] = \mu$. Then, with probability 1,

$$\frac{X_1 + X_2 + ... + X_n}{n} \to \mu \qquad \text{as } n \to \infty$$

8.6 Central Limit Theorem

The Central Limit Theorem provides a simple method to compute approximate probabilities of sums of independent random variables.

Central Limit Theorem

Let $X_1, X_2, ..., X_n$ be a sequence of independent and indentically distributed random variables, each with mean μ and variance σ^2. Then the distribution

$$\frac{X_1, ..., X_n - n\mu}{\sigma\sqrt{n}}$$

tends to the standard normal as $n \to \infty$. So

$$P\left\{\frac{X_1, ..., X_n - n\mu}{\sigma\sqrt{n}} \leq a\right\} \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-x^2/2}\ dx \qquad \text{as } n \to \infty$$

9 Markov Chains

Markov Chains model a system of N states where there is a probability of moving from one state to another at every click of a clock. The probability of moving to state i to state j depend entirely on i and not the number of clock "clicks" passed.

$$P_{i,j} = P(\text{system is in state } j \text{ at time } n+1 | \text{ system in state } i \text{ at time } n)$$

Such probabilities are written in a *transition matrix.*

$$P = \begin{bmatrix} P_{0,0} & P_{0,1} & P_{0,2} \\ P_{1,0} & P_{1,1} & P_{1,2} \\ P_{2,0} & P_{2,1} & P_{2,2} \end{bmatrix}$$

where $P_{i,j}$ is the probability of transitioning from state i to state j at this time. The n^{th} probability vector $\pi^{(n)}$ is a vector of the probabilities that we are in state i at time n, where

$$\pi^{(n)} = \left(\pi_0^{(n)}, \pi_1^{(n)}, \pi_2^{(n)}\right)$$

Properties of the n^{th} probability vector

$$\pi^{(n+1)} = \pi^{(n)} P$$

$$\pi^{(n)} = \pi^{(0)} P^n$$

For large n, if $\lim_{n\to\infty} \pi^{(n)}$ exists and is independent of π^0 then

$$\pi = \lim_{n\to\infty} \pi^{(n)}$$

is the **steady state probability vector**. It can be found by:

1. Solving for $\pi = \pi P$
2. Use $\pi_0 + ... + \pi_{N-1} = 1$

assuming it exists (ergodic). It **may not** if

1. Periodic Behaviour where the chain goes back and forth
2. Traps where a chain gets stuck in one state or another, in which case it depends on $\pi^{(0)}$

Ergodic Markov Chain A markov chain is ergodic iff there exists $n \in \mathbb{N}^+$ such that P^n has no zero entries. It is aperiodic and irreducible. If so, then the **steady state probability vector** exists and π is independent of $\pi^{(0)}$.

10 Markov Chains in Continuous Time

If, instead of a system changing at discrete times, it can change at any time, and the time between transitions are *exponentially distributed*, then the transitional rate probabilities are given by a **Poisson Process**. Here we focus on calculate the resulting steady state probability vector.

10.1 Poisson Process

We can model things like the number of buses $\tilde{N}(t)$ arriving in a time period $[t, t + \sigma t]$. This Poisson Process has some properties:

Properties of a Poisson Process $\tilde{N}(t)$

1. For any fixed t, $\tilde{N}(t)$ is a discrete random variable
2. $\tilde{N}(0) = 0$, so counting starts at time 0
3. # of events in disjoint intervals are independent
4. $\tilde{N}(t+h) - \tilde{N}(t) = \#$ of events in $[t, t+h]$ for $h \to 0$
5. $P(\tilde{N}(h) = 1) = P(\text{event occurs in } [t, t+h]) = \lambda h + E(h)\text{as } \frac{E(h)}{h} \to 0, h \to 0 \qquad E(h)$ is small even in relation to h
6. $\frac{1}{h} P(\tilde{N}(h) \geq 2) \to 0 \qquad$ as $h \to 0$

If the above are satisfied then

$$P(\tilde{N}(t) = k) = \frac{(\lambda t)^k}{k} e^{-\lambda t} \qquad k = 0, 1, 2, ...$$

The number of births in interval $(0, t) \approx Poisson(\lambda t)$

10.2 Birth Death Process

A special case of continuous time Markov Chains where within some small time period δt transitions can only increase the state variable by 1.

Birth Rates $\lambda_{i,i+1} = b_i$
Death Rates $\lambda_{i,i-1} = d_i$
otherwise $\lambda_{i,j} = 0$
If each b_i and d_i is non-zero, and N is finite, then the steady state probability vector exists and it is independent of $\pi^{(0)}$

Calculating Steady State Probability Vector for a Birth-Death Process

In general,

$$b_j \pi_j = d_{j+1} \pi_{j+1} \qquad j = 0, 1, 2..., N-1$$

1. $\pi_j = \frac{b_0 b_1 ... b_{j-1}}{d_1 ... d_j} \pi_0 \qquad j = 1, 2, ..., N-1$
2. $1 = \pi_0 + \pi_1 + ... + \pi_{N-1}$

M/M/S Queues Birth-Death Processes with particular choice of birth rates and death rates and in infinite number of possible states (length of the queue).

1. Customers arrive as described by Poisson Process with rate λ.
2. $S = \#$ of servers
3. Service time of Server (time to service a customer) exponentially distributed with mean $\frac{1}{\mu}$
4. State 'j' mean j customers are in queue, $j = 0, 1, 2...$
5. The actual state: discrete $J \in \{0, 1, ..., N-1\}$
6. $b_j = \lambda \qquad j = 1, 2, ..., S$
7. $d_j = \begin{cases} j\mu, & j = 1, 2, ..., S \\ S\mu, & j \geq S \end{cases}$

If $\lambda < S_\mu$ then a steady state distribution $\pi = (\lambda_0, \lambda_1, ...)$ exists, and is independent of $\pi^{(0)}$.
π_j is the Probability Mass Function of state J: $P($J customers in queue$\}$

Following page describes the steady state probability vector of **M/M/1**, the mean queue length and an example.

Steady state probability vector

$$\text{(A)}\ \pi_j = \frac{b_0 \cdots b_{j-1}}{d_1 \cdots d_j}\pi_0 = \left(\frac{\lambda}{\mu}\right)^j \pi_0, \qquad j = 0, 1, 2, \cdots$$

$$\text{(B)}\ 1 = \pi_0 + \pi_1 + \pi_2 + \cdots = \pi_0\Big[1 + \frac{\lambda}{\mu} + (\frac{\lambda}{\mu})^2 + \cdots\Big] = \pi_0 \sum_{j=0}^{\infty}\left(\frac{\lambda}{\mu}\right)^j = \pi_0 \frac{1}{1-\frac{\lambda}{\mu}}$$

$$\text{if } \lambda < \mu \quad \Rightarrow \quad \pi_0 = \frac{\mu - 1}{\mu} = 1 - \frac{\lambda}{\mu} \qquad \Rightarrow \quad \pi_j = \left(1 - \frac{\lambda}{\mu}\right)\left(\frac{\lambda}{\mu}\right)^j, \qquad j = 0, 1, \cdots$$

$$\mathbb{E}[J] := \sum_{j=0}^{\infty} j\pi_j = \left(1 - \frac{\lambda}{\mu}\right)\sum_{j=1}^{\infty} j\left(\frac{\lambda}{\mu}\right)^j = \left(\frac{\lambda}{\mu}\right)\left(1 - \frac{\lambda}{\mu}\right)\sum_{j=1}^{\infty} j\left(\frac{\lambda}{\mu}\right)^{j-1} = \ ?$$

Recall: $\displaystyle\sum_{j=0}^{\infty} x^j = \frac{1}{1-x}, \quad 0 < x < 1$

$$\overset{\text{differentiate}}{\Rightarrow} \quad \sum_{j=1}^{\infty} j x^{j-1} = \frac{1}{(1-x)^2}$$

Integrals, sums and limits - ShareLaTeX, Online LaTeX Editor

Hence,

$$\mathbb{E}[J] = \frac{\lambda}{\mu}\left(1 - \frac{\lambda}{\mu}\right)\frac{1}{\left(1-\frac{\lambda}{\mu}\right)^2} = \frac{\frac{\lambda}{\mu}}{1-\frac{\lambda}{\mu}} = \frac{\lambda}{\mu - \lambda}$$

Example An M/M/1 queue has customers arriving at 9/min w/ service rate 10/min.

(1) What is the mean queue length in the steady-state?

Solution: $\mathbb{E}[J] = \dfrac{\lambda}{\mu - \lambda} = \dfrac{9}{10-9} = 9$

(2) If there are 2 servers instead of 1 (M/M/2 queue)
What is the probability that no customers are in the queue in the steady-state?

Solution: $b_j = \lambda = 9, \quad j = 0, 1, 2, \cdots \qquad d_1 = \mu = 10, \quad d_j = 2\mu = 20, \quad j \geq 2$

(A) $\pi_j = \dfrac{b_0 b_1 \cdots b_{j-1}}{d_1 d_2 \cdots d_j}\pi_0$

$$\pi_1 = \frac{b_0}{d_1}\pi_0 = \frac{\lambda}{\mu}\pi_0 = \frac{9}{10}\pi_0 \qquad \pi_j = \frac{b_0\, b_1 \cdots b_{j-1}}{d_1\ d_2 \cdots d_j}\pi_0 = \frac{\lambda}{\mu}\left(\frac{\lambda}{2\mu}\right)^{j-1}\pi_0 = \frac{1}{2^{j-1}}\left(\frac{\lambda}{\mu}\right)^j \pi_0$$

(B) $$1 = \pi_0 + \pi_1 + \cdots = \pi_0\Big[1 + \frac{\lambda}{\mu} + 2\sum_{j=2}^{\infty}\left(\frac{\lambda}{2\mu}\right)^j\Big] = \pi_0\left[1 + \frac{\lambda}{\mu} + 2\frac{\left(\frac{\lambda}{2\mu}\right)^2}{1-\frac{\lambda}{2\mu}}\right]$$

$$\pi_0 = \frac{1}{1 + \frac{9}{10} + \frac{1}{10}\frac{9^2}{20-9}} = \frac{10}{19 - \frac{81}{11}} \approx 0.508$$

20

11 Entropy Equations

The convention now is that log is in fact log_2.

Entropy of X

$$H(X) = -\sum_k p_X(x_k) log p_X(x_k)$$

Surprise

$$S(X) = -log(p)$$

$$S(X = x_k) = -log_{p_X}(x_k)$$

Average Uncertainty

$$H(X,Y) = -\sum_j \sum_k p_{X,Y}(x_j, x_k) log p_{X,Y}(x_j, y_k)$$

$$H(X,Y) = H(Y) + H_Y(X) = H(X) + H_X(Y)$$

If X and Y are independent

$$H(X,Y) = H(Y) + H(X)$$

Uncertainty of X given Y

$$H_{Y=y_k}(X) = -\sum_j p_{X|(Y=y_k)}(x_j) log p_{X|(Y=y_k)}(x_j)$$

Conditional Entropy

$$H_Y(X) = \sum_k H_{(Y=y_k)}(X) p_Y(y_k)$$

$$H_Y(X) = \sum_x \sum_y P(x,y) log \frac{P_Y(y)}{p(x,y)}$$

Desired properties for $\mathcal{S}$:

(A) $\mathcal{S}(1) = 0 \neq \mathcal{S}(0)$, which is undefined or $+\infty$

(B) $\mathcal{S}$ decreases: $p < q \Rightarrow \mathcal{S}(q) < \mathcal{S}(p)$

(C) $\mathcal{S}(pq) = \mathcal{S}(p) + \mathcal{S}(q)$

Theorem 3.1. *If $\mathcal{S}$ is continuous and conditions* (A)-(C) *above are satisfied, then there is a constant $\mathscr{C} > 0$ such that for all $p \in [0,1]$, $\mathcal{S}(p) = -\mathscr{C}\log(p)$.*

TABLE 5.1: AREA $\Phi(x)$ UNDER THE STANDARD NORMAL CURVE TO THE LEFT OF X

X	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998