Chapter 1

Conditional probability

1.1 Total Probability theorem

P(B) $= P(A_1)P(B|A_1) + ... + P(A_n)P(B|A_n)$

1.2 Baye's Rule

A Prior probability is Before we have any evidence. A Posterier probability is after the evidence.P(B|A) is the probability of B after event A has happened. This doesn't mean that P(A)=1. Also, this doesn't change the prior probability.

Baye's Rule in the discrete case
$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$

Baye's Rule in the continuos case
$f_{X|Y}(x|y) = \frac{f_{Y|X}(y|x)f_X(x)}{f_Y(y)}$

Baye's rule also works for a mix of Discrete and Continuous random variables

1.3 Joint and Marginal PDF

The marginal PDFs of X and Y can be calculated from the joint PDF, using the formulas
$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)$

Chapter 2

Continuous Random Variables

2.1 Expected Value

Expected value of an event is = Probability of that event * Value of the outcome. If Probability of winning a lottery is .01 and the amount is Dollar 100. Then the Expected winnings are 100*.01 = 1 $E[X] = \int_{-\infty}^{\infty} xf(x)dx$
$E[X] = \int_{0}^{\infty} P(X > y)dy - \int_{-\infty}^{0} P(X < y)dy$

Expected value of a function of a Random Variable
$E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)dx$
Linearity of Expectations
E[aX + b] = aE[X] + b

2.2 Variance

Variance is the expected value of the function $g(X) = (X - E[X])^2$.
Var(X) = $E[(X - E[X])^2]$
Var(X) = $E[X^2] - E[X]^2$
Var(aX + b) = $a^2Var(X)$

Chapter 3

Further topics on RVs

3.1 Derived Distributions

We know X, and we want Y = g(X). A general method of solving this type of problem is
1.Find the CDF of Y : $F_y = P(Y \leq y)$
2.Substitute Y with X using the g(.)
3.Rewrite $F_y as F_x$
4.Differentiate

3.2 Convolution : Sum of two Random variables

X + Y = Z
When X and Y are independent E[Z] = E[X] + E[Y]

$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)$
i.e. Take all the X,Y pairs whose sum is Z. This formula is valid for independent RVs

Sum of two normal distributions
A = X + Y, where X and Y are independent normal. Then A is also normal with E[A] = E[X] + E[Y] and var(A) = var(X) + var(Y). Do not use this formula when X and Y are not independent.

3.3 Conditional expectations

E[X|Y=y] is the expected value of X, when Y = y. Due to the condition, E[X|Y] is a function of Y.

Law of iterated Expectations: E[E[X|Y]] = E[X]

3.4 Covariance and corelation

Covariance
Cov(X,Y) = E[(X-E[X])(Y-E[Y])]
Cov(X,Y) = E[XY] - E[X]E[Y]

Properties of Covariance
If X and Y are independent then Cov(X,Y)=0. The converse is not always true.

Cov(X,X) = Var(X)
Cov(aX + b , Y) = aCov(X,Y)
Cov(X , Y + Z) = Cov(X,Y) + Cov(X,Z)

Variance of a sum of RVs

$Var(\sum_{i=1}^{n} X_i) = \sum_{j=1}^{n} Var(X_j) + \sum_{(i,j),i\neq j} Cov(X_i, X_j)$

Variance of sums = Sum of Variances + Sum of all the combinations of the covariances. If they are **independent**, then the covariance = 0 and the variances of the sum = sum of the variances.
Co-relation co-efficient, $\rho(X,Y) = \frac{Cov(X,Y)}{\sigma_x \sigma_y}$

3.5 The Conditional Variance

Law of total variance
var(X) = E[**var(X|Y)**] + var(**E[X|Y]**)
The bold stuff are a random variable of Y. So, they have an expected value as well as a variance
How are these useful? I know how to deal with the sum of Random Variables. But what about the sum of random number of random variables. i.e. that the number of random variables is also a random variable.

Example. I go shopping.
N: Number of shops visited
X_i: Money spent in each shop
$Y = X_1 + ...X_N$ *Note that N is itself a RV*
Find the E[Y]. i.e. The expected value of the total money spent I know how to deal with this if I know the number of shops that were visited. So, I will condition on that.

$$
\begin{aligned}
E[Y|N = n] &= E[X_1 + ... + X_n|N = n] \\
&= E[X_1 + ... + X_n] \\
&= nE[X]
\end{aligned}
$$

By the **Total Expectation Theorem**

$$\begin{aligned} E[Y] &= \sum_n p_n E[Y|N=n] \\ &= \sum_n p_n n E[X] \\ &= E[X] \sum_n p_n n \\ &= E[X]E[N] \end{aligned}$$

By the **Law of Iterated Expectations**

$$\begin{aligned} E[Y] &= E[E[Y|N]] \\ E[Y] &= E[NE[X]] \\ &= E[X]E[N] \end{aligned}$$

This makes intuitive sense. The amount of money spent is equal to the amount spent in each store multiplied by the number of stores.

Now I find the Variance of a random number of random variables

By the **Law of Total Variance**

var(Y) = E[var(Y|N)] + var(E[Y|N])
Now, I find the two terms.

$$\begin{aligned} E[var[Y|N]] &= E[Nvar(X)] \\ &= var(X)E[N] \\ var(E[Y|N]) &= var(NE[X]) \\ &= (E[X])^2 var(N) \\ var(Y) &= var(X)E[N] + (E[X])^2 var(N) \end{aligned}$$

Chapter 4

Bayesian Statistical Inference

in BI, to find an unknown quantity/model we assume it to be a random variable θ. This random variable has a prior probability distribution $p_\theta(\theta)$. We observe some data x, and use Baye's rule to derive a posterior probability distribution $p_{\theta|X}(\theta|x)$. This captures all the information that x can give about θ.

Summary of Bayesian inference
1. We start with a prior distribution p_θ or f_θ for the unknown RV.
2. We have a model $p_{X|\theta}$ or $f_{X|\theta}$. i.e. How is X like, when θ is true. X is the thing that we can observe.
3. For the posterior distribution of θ use Baye's rule.

Chapter 5

Problem solving tips

1. Deductibles : Suppose my deductible is \$100. Then if the total bill is less than 100, the insurance company will pay nothing. If the bill is\$200, the the company will pay \$200 - \$100 = \$100.

2. $E[X^2]$ can be easily found using the $Var(X) = E[X^2] - E[X]^2$ formula. Otherwise, use the formula $\sum_x x.p(x)$ or $\int x.f(x)$

3. Many complex problems comes down to Binomial in the end. Be careful and do not miss the $\binom{n}{k}$ term.

4. While changing the variables in an integration. I often mess up and forget about changing the limits. Do, not be lazy and change the limits when you change the variable.

5. Only an IID distribution can be modelled as Binomial. i.e. All the trials should be independent. In some problems, like the unknown bias of a coin. The trials are conditionally independent. Using this condition, the problem can be solved using the law of iterated expectation.

6. **Wishful thinking**. If knowing a random variable helps me solve the problem. Then try to condition on that random variable and then think about the problem. Example, if I need to find the probability of $P(X < Y)$ where X and Y are independent Random Variables. Then if I can find this $P(X < y|Y = y)$. Then, $P(X < Y) = P(X < y|Y = y)f_Y(y)$. Add this for all values of Y. Or integrate in case of continuous RV. $P(X < Y) = i.e. \int_Y P(X < y|Y = y)f_Y(y)$.

7. **Maximum values** If Y = $max(X_1, X_2, X_3)$. What is the $P(Y > y)$. If at least one of the $X_i > y$. Then Y will also be greater than y. So, this can be modelled as $P(X_1 \cup X_2 \cup X_3)$. Look at the next point to find P(Y¡y).

8. Another similar problem is in page 239. I need to find out the Expected value of the largest claim made. The PDF of the claim f_X is given.

Find the CDF F_X. Now, $F_Y(y) = P(Y < y) = P(max(X_1, X_2, X_3) < y)$. Suppose, the maximum value of Y is 10, then any of the X_i can not be greater than 10. So, this means $P(X_1 < y) \cap P(X_2 < y) \cap P(X_3 < y)$. This gives me the CDF of largest claim made. Which is give me the PDF of the largest claim made and then I can find the expected value. Actually, the same method can be used to solve the previous problem.

9. For problems involving deductibles create a new RV for the actual payment. Example, if the PDF for loss is f(x) = $\begin{cases} \frac{1}{x^2}\, for\, x > 1 \\ 0,\ otherwise \end{cases}$ And the deductible is 2. Then define Y, where Y is the retained losses of the policyholder. If the actual loss is between 1 and 2, then the policyholder has to pay for everything. and if the actual loss is greater than 2, then the insurance company will pay the actuals minus the deductible Y = $\begin{cases} X\, for \quad 1 < X < 2 \\ X - 2\, for\, X > 2 \end{cases}$

   For the insurance company, Define another random variable Z, representing what the company pays. The company pays nothing if the X is less than the deductible. Z = $\begin{cases} 0\, for 1 < X < 2 \\ X - 2\, for X > 2. \end{cases}$

   Wrong method E[loss] - deductible doesn't give the E[payment]. It doesn't account for the case when the loss is less than the deductible and the payment (by the insurance company) is zero.

10. PDFs with the term $|x|$ can often be solved faster by using symmetry. If the function is symmetric around a point, then that point is the expected value. If a part of the function is symmetric around zero. Then that part can be ignored. Convince yourself of symmetry by drawing the function.

11. **Percentile** Given PDF, what is the 95th percentile? It is asking for the value of X = x for which F(x) = .95 or P(X ≤ x)=.95.

12. How to find the **median**? Find the CDF, F(X = x). F(x) = .5

13. **Survival function** is the complement of the CDF. i.e. $P(X < x)$

14. The concept of derived distributions is sometimes useful in calculating payments. When the relationship between the RV of payment and RV of loss is given. For example.
    The time, T, that a manufacturing system is out of operation has CDF. F(t) = $\begin{cases} 1 - (\frac{2}{t})^2, t > 2 \\ 0, otherwise. \end{cases}$ The resulting cost to the company is $Y = T^2$. Determine the density function of Y, for $y > 4$.

    (a) Write F$_Y(y)$ =P(Y ≤ y)

(b) Replace Y with T using g(.).
$P(T^2 \leq y) = P(T \leq \sqrt{y}) = 1 - (\frac{2}{\sqrt{y}}) = 1 - \frac{4}{y}^2$

(c) $f(y) = F'(y) = \frac{4}{y^2}$

15. It is always useful to write the payment random variable before doing any calculations. For example $f(y) = 2y^{-3}$ for $y > 1$. Where y is the loss value. And it is given that the payment is capped at 10. So, The RV for payment will be, Y= $\begin{cases} 0, Y \leq 1 \\ Y, for 1 < Y < 10 \\ 10, for Y > 10 \end{cases}$. Now that I've divided this up into cases, I can easily calculate the Expected Value.

Chapter 6

Integration

1. $\int_0^\infty x^t e^- t dt = t!$