*P*(*A* ∩ *B*) = *P*((*Ac* ∪ *Bc*)*c*)

*P*(*A* ∩ *B*) = *P*(*A*) + *P*(*B*) − *P*(*A* ∪ *B*)

*P*(*A* ∪ *B*) = *P*(*A*) + *P*(*B*) − *P*(*A* ∩ *B*)

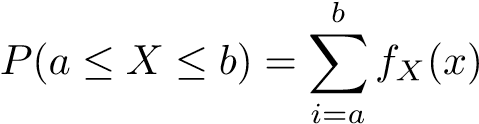
*P*(*A*) = 1 − *P*(*Ac*)

1. ≤ *P*(*A*) ≤ 1
2. **Equations from this section**: Shape

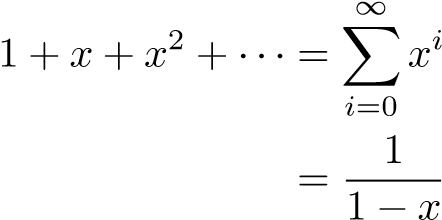
   Description automatically generated with medium confidence

* **Probability Intervals**:

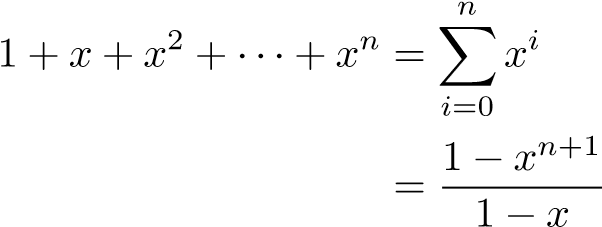
If *X* is a DRV with pmf *fX*(*x*), then:



* **Geometric Series**
  + *Infinite Geometric series*:

*,* for 0 *< x <* 1

* + *Finite Geometric Series*:

*,* for 0 *< x <* 1

* Law of total probability applied to DRVs

If *X* is a DRV with pmf *fX*(*x*) and cdf *FX*(*x*), then

*P*(−∞ *< X <* ∞) = X *fX*(*i*) = 1

*i*=−∞

*FX*(∞) = *P*(*X* ≤ ∞) = X *fX*(*i*) = 1

* **Common Discrete Distributions**
  + *Bernoulli Distribution* {bern(*p*)}:

The Bernoulli distribution represents flipping an unfair coin that has probability *p* of coming up heads, and counting a heads as a 1, and tails as a 0.

If *X* ∼ bern(*p*)

*fX*(*x*) = (*p*)*x*(1 − *p*)1−*x*

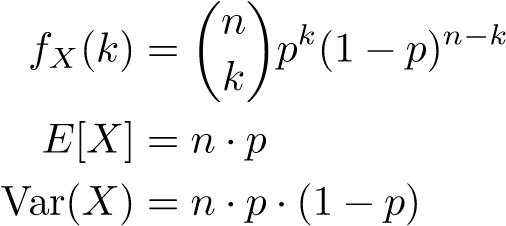
For *x* ∈ {0*,*1}

*E*[*X*] = *p*

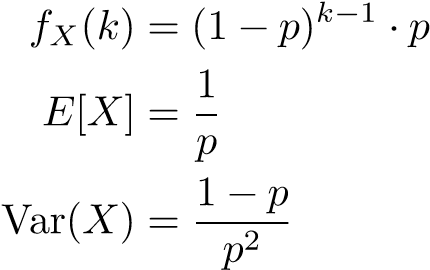
Var(*X*) = *p*(1 − *p*)

* + *Binomial Distribution* {bin(*n,p*)}

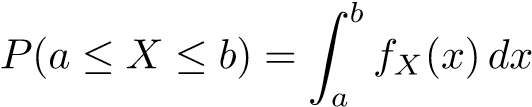
The binomial distribution is a sum of Bernoulli random trials If *X* ∼ bin(n,p), then:

*,* for *k* = 0*,*1*,*2*,*··· *,n*

* + *Geometric Distribution* {Geo(*p*)} A geometric distribution can be described as the number of Bernoulli trials required to get a success If *X* ∼ Geo(p), then:



* The major difference between a DRV and a continuous random variable (CRV) is that we switch from interpreting probabilities as sums of finite (or countably infinite) points, to areas under the curve.
  + For example: If *X* is a CRV, then we can always say *P*(*X* = *c*) = 0, where *c* is any constant
* **Probability Density Function** If *X* is a CRV, then it has a probability density function (pdf) *fX*(*x*), such that



* When talking about CRV’s, we note that the probability of a CRV being exactly equal to any particular number is 0

This also means that

*P*(*a* ≤ *X* ≤ *b*) = *P*(*X* = *a*) + *P*(*a < X < b*) + *P*(*X* = *b*)

= 0 + *P*(*a < X < b*) + 0

=⇒ *P*(*a* ≤ *X* ≤ *b*) = *P*(*a < X < b*)

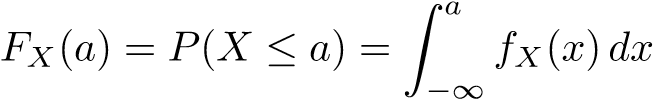
So we can exchange inclusive inequalities (≤*,*≥) with exclusive inequalities

(*<,>*) as we want when we are talking about CRVs

* **Cumulative Distribution Function**

One thing that DRVs and CRVs have in common is that they both have a cdf

So if *X* is a CRV with pdf *fX*(*x*), then *X* also has a cdf *FX*(*x*), such that:



* Law of total probability applied to CRVs

In the same vein as DRVs, if *X* is a CRV with pdf *fX*(*x*) and cdf *FX*(*x*), then

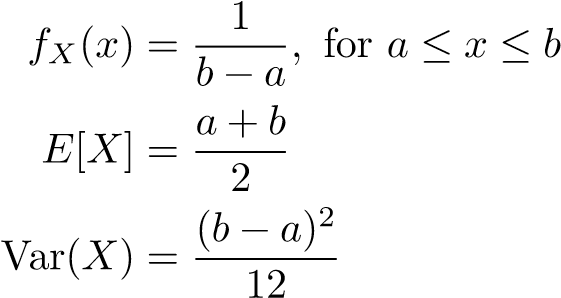
Z ∞ *P*(−∞ *< X <* ∞) = *fX*(*x*)*dx* = 1

−∞ Z ∞ *FX*(∞) = *P*(*X* ≤ ∞) = *fX*(*x*)*dx* = 1

−∞

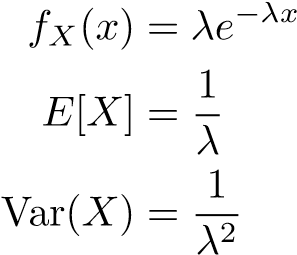
* **Common Continuous Distributions**:
  + *Uniform Distribution* {Unif(*a,b*)}

This distribution describes the event where the random variable is some value inside the interval [*a,b*], and it can take on any of the values in that interval with the same probability If *X* ∼ Unif(*a,b*), then

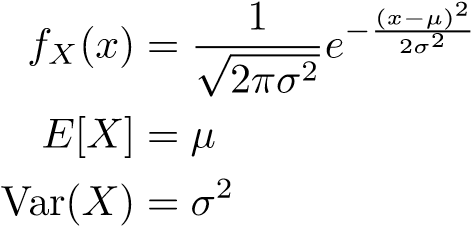


* + *Exponential Distribution* {Exp(*λ*)}

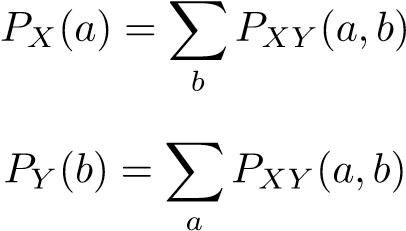
If *X* ∼ Exp(*λ*) for *λ >* 0, then



* + *Normal Distribution* {N(*µ,σ*2)} If *X* ∼ N(*µ,σ*2), then



1. Discrete R.V:
   1. Joint pmf: *P*(*X* = *a,Y* = *b*) or *PXY* (*a,b*).
   2. Marginalization: *PXY* − − *> PX* and *PY* ?



(sum rows and columns in the table). **See quiz 5, problem 2**

* 1. Joint cdf: *FXY* (*a,b*) = *P*(*X* ≤ *a,Y* ≤ *b*)
  2. *FXY* − − *> FX,FY* ?

*FX*(*a*) = *P*(*X* ≤ *a*) = *FXY* (*a,*∞) (Let *b* = ∞)

*FY* (*b*) = *P*(*Y* ≤ *b*) = *FXY* (∞*,b*) (Let *a* = ∞) **See quiz 5, problem 1**

1. Continuous R.V.
2. Joint pdf: *f*(*x,y*).
3. Marginalization: *f*(*x,y*) − − *> f*(*x*) and *f*(*y*)?

*f*(*x*) = *f*(*x,y*)*dy*

*f*(*y*) = *f*(*x,y*)*dx*

**See hw5, problem 6**

1. Given the joint pdf f(x, y), how to compute probability?

(e.g. find *P*(*a* ≤ *X* ≤ *b,c* ≤ *Y* ≤ *d*))

Z Z

*P*(*a* ≤ *X* ≤ *b,c* ≤ *Y* ≤ *d*) = *f*(*x,y*)*dydx*

Note that the ranges of variables are not filled in these equations. It is the critical part of solving problems in continuous R.V.

3) Independence:

If X, Y are independent,

1. Discrete R.V.: *P*(*X* = *a,Y* = *b*) = *P*(*X* = *a*)*P*(*Y* = *b*)
2. Continuous R.V.: *f*(*x,y*) = *f*(*x*)*f*(*y*) (*f*(*x*) and *f*(*y*) can be found by marginalization)

* 1. *Cov*(*X,Y* ) = *E*[(*X* − *E*(*X*))(*Y* − *E*(*Y* ))]
* 2. *Cov*(*X,X*) = *V ar*(*X*)
* 3. *Cov*(*X,Y* ) = *Cov*(*Y,X*)
* 4. *Cov*(*X,c*) = *E*[(*X* − *E*(*X*))(*c* − *E*(*c*))] = 0
* 5. *Cov*(*X,Y* + *Z*) = *Cov*(*X,Y* ) + *Cov*(*X* + *Z*)

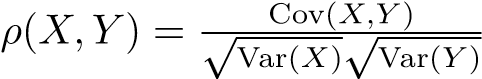
*Cov*(*X* + *Y,Z* + *K*) = *Cov*(*X,Z*) + *Cov*(*X,K*) + *Cov*(*Y,Z*) + *Cov*(*Y,K*)

(Bilinearily property)

* 6. *Cov*(*X,Y* ) = *E*(*XY* ) − *E*(*X*)*E*(*Y* )
* 7. If two R.V. X and Y are independent Cov(*X,Y* ) = *E*(*XY* )−*E*(*X*)*E*(*Y* ) = *E*(*X*)*E*(*Y* ) − *E*(*X*)*E*(*Y* ) = 0

But given *Cov*(*X,Y* ) = 0, we cannot say X and Y are independent. (One way)

**See hw6 problems**

. We have −1 ≤ *ρ* ≤ 1

1). Height of a bin:

