Probability and Statistics

သတိ: စာကိုမကျက်ပါနှင့်။
နားလည်အောင်ဖတ်ပြီးစဉ်းစားပါ။

Will you Gamble or Risk?

Russian Roulette လို့ ခေါ်တဲ့ Game တစ်ခု ရှိပါသည်။ Revolver (ခြောက်လုံးပြူး) ထဲကို ကျည်ဆန် တစ်တောင့် ထည့်ထားမည်။ ပြီးရင် ဆုံလည်ကို လှည့်လိုက်ပြီး ကိုယ့်ခေါင်းကို ကိုယ် ပြန်ပစ်မည် ဆိုပါတော့။ ကျည်မထွက်လာရင် ဒေါ်လာ 1 သန်းရမယ်ဆိုပါတော့။ ကျည်ထွက်လာရင်တော့...

ကျွန်တော်တို့ မေးခွန်းတစ်ခုနဲ့ စရအောင်ပါ။
ဘယ်ဟာကို ရွေးမလဲ။

ဒီ 2 ခုမှာ ဆိုရင် ဘာကို ရွေးမလဲ။ ဒီ 2 ခုဘာကွာတာလဲ ဆိုရင် ဘယ်လို ဖြေမလဲ။

အမှန်တော့ ပထမ တစ်ခုမှာ ကျွန်တော်တို့က ဘာကို သိလဲလို့ဆိုရင် Probability ကို ကြိုသိပါသည်။ ခင်ဗျား သေဖို့သည် 6 ပုံ တစ်ပုံ ဖြစ်ပြီး ဒေါ်လာ 1 သန်းရဖို့ သည် 5 ပုံ 6 ပုံ ဖြစ်သည်။ ပိုပြီး အရေးကြီးတာ ဒီ Probability ကို ခင်ဗျား ပြင်လို့မရ။ Fixed Probability ဖြစ်သည်။ ဘယ်လိုပဲ ကြိုးစားကြိုးစား သေဖို့သည် 6 ပုံ တစ်ပုံ ဖြစ်ပြီး ဒေါ်လာ 1 သန်းရဖို့ သည် 5 ပုံ 6 ပုံ ဖြစ်သည်။

ဒုတိယ တစ်ခုက ကျွန်တော်တို့ Probability ကို ကြိုမသိပါဘူး။ ခင်ဗျား ပြုတ်ကျဖို့ Probability ဘယ်လောက်ရှိလဲ ကြိုမသိပါဘူး။ ပိုအရေးကြီးတာက ဒီ Probability ကို ခင်ဗျား ပြင်လို့ ရသည်။ လေ့ကျင့်ဖို့ 6 လ ရှိသည့် အတွက် ခင်ဗျား ကြိုးစားရင် ကြိုးစားသလို ပြုတ်ကျဖို့ နည်းသွားမှာ ဖြစ်သည်။

ပထမ အမျိုးအစားကို Gamble (အလောင်းအစား) ဟုခေါ်ပြီး ဒုတိယ အမျိုးအစားကို Risk (စွန့်စားမှု) ဟုခေါ်သည်။

8 ထပ်တိုက် 2 ခုကြားက ပေ 40 လောက်ဝေးသည်။ တိုက် 2 ခုရဲ့ ခေါင်မိုးများကို လက် 3 လုံးလောက် ရှိတဲ့ သံမဏိချောင်းနဲ့ ဆက်ထားသည်။ တစ်ဖက်ခေါင်မိုးမှ အခြားတစ်ဖက်သို့ သံမဏိချောင်းပေါ်ကနေ ကူးသွားရမယ် ဆိုပါတော့။ ပြီးရင် ဒါကို လေ့ကျင့်ဖို့ 6 လ အချိန်ပေးမည် ဆိုပါတော့။ ပြုတ်မကျဘဲ ရောက်အောင် ကူးသွားနိုင်ရင် ဒေါ်လာ 1 သန်းရမယ်ဆိုပါတော့။ ပြုတ်ကျသွားရင်တော့...

Non Deterministic Causality

- Risk ပဲ ဖြစ်ဖြစ်၊ Gamble ပဲ ဖြစ်ဖြစ် Result ကို ကျွန်တော်တို့ ကြိုမသိပါဘူး။ ဒိဋ္ဌေဒိဋ္ဌ မပြောနိုင်ပါဘူး။
- ဒီလို ဒိဋ္ဌေဒိဋ္ဌ မပြောနိုင်တဲ့ ကြောင်းကျိုး ဆက်စပ်မှုကို Non Deterministic Causality ဟု ခေါ်ပါသည်။
- ကျွန်တော်တို့ Calculus သည် Deterministic Causality နှင့် Deterministic Change ကို လေ့လာသော သင်္ချာပညာ ဖြစ်သည် လို့ ပြောခဲ့ပါသည်။
- ဒီတစ်ခါတော့ Non Deterministic Causality ကို လေ့လာသော သင်္ချာပညာ အကြောင်းကို ဆွေးနွေးကြပါမည်။
- ဒါကို Mathematics of Chance လို့ ခေါ်မည်ဆိုလည်း ခေါ်နိုင်ပါသည်။
- ဒီ သင်္ချာပညာ သည် Probability နှင့် Statistics ဖြစ်သည်။

Population and Sample

- Population သည် State Space တစ်ခု ဖြစ်သည်။ ဆိုလိုသည်က A Set of All (Complete) States ကို ဆိုလိုတာ ဖြစ်သည်။ A Set of All Numbers ဆိုရင်လဲ ဂဏန်းအားလုံးကို ဆိုလိုတာ ဖြစ်သည်။
- မြန်မာပြည်၏ လူဦးရေဆိုလျှင် ဗိုက်ထဲက ကလေးမှ စ၍ သေလုနီးပါး လူနာများ အထိ ပါသည်။
- အမှန်တော့ Population တစ်ခုလုံးကို ဘယ်သူမှ မသိပါ။ ဒါသည် Theoretical Assumption တစ်ခု ဖြစ်သည်။
- Coin တစ်ခုကို ခေါင်းပန်း လှန်ရင်တောင် ဒေါင်လိုက်ကျ နိုင်သေးသည်။ ဖြစ်ဖို့တော့ အင်မတန်နည်းသည်၊ ဒါပေ့မယ် ဖြစ်နိုင်သည်။
- ထို့ကြောင့် မြန်မာပြည်၏ လူဦးရေဆိုလျှင် ရှိသမျှ လူအားလုံးမဟုတ်တော့ဘဲ စစ်တမ်းကောက်နိုင်သော လူဦးရေကို သာဆိုလိုပါသည်။
- ထို့အတူ ခေါင်းပန်းလှန်ရင်လဲ ခေါင်းနှင့် ပန်းကိုသာ စဉ်းစားသည်။ ဒေါင်ကျခြင်းကို ထည့်မပြောတော့။
- ဒါကို Sample Space ဟုခေါ်သည်။ Sample Space သည် A Set of All Possible (Observable) States ကို ပြောတာဖြစ်သည်။

Sampling

- Population တစ်ခုလုံးကို ဘယ်သူမှ မသိနိုင်သည့် အတွက် Population မှ Sample များကို ရယူခြင်းသည် အရေးကြီးပါသည်။ ဒါကို Sampling (စစ်တမ်းကောက်ခြင်း) ဟုခေါ်သည်။

- Sampling လုပ်သော နည်းမျိုးစုံရှိပါသည်။

- Random Sampling သည် Population မှ Sample များကို Randomly စစ်တမ်းကောက်ခြင်း ဖြစ်သည်။ ဥပမာ၊ အမျိုးသား၊ အမျိုးသမီးဦးရေကို သိချင်ပါက လမ်းပေါ်မှာ လူများကို Randomly စစ်တမ်းကောက်ခြင်း ဖြစ်သည်။

- Random Sampling မှာ အားနည်းချက်က Bias ရှိနိုင်ပါသည်။ ဥပမာ၊ ကော်ဖီသောက်သူများကို စစ်တမ်းကောက်မည်ဆိုပါတော့။ ဒါဆိုရင် လူဦးရေ 1000 ရှိသော မြို့လေးတစ်မြို့မှာ လူ 100 စစ်တမ်းကောက်သည် ဆိုပါတော့။ သို့သော် လူဦးရေ 100,000 မြို့ကြီးတစ်မြို့မှာ လူ 100 ပဲ စစ်တမ်းကောက်လို့ မရတော့။ အနည်းဆုံး လူ 10,000 ကို စစ်တမ်းကောက်မှ ရသည်။ ဒါကို Systematic Sampling ဟု ခေါ်သည်။

- စစ်တမ်းကောက်ခြင်း (Sampling) လုပ်ခြင်းမှ Data (Sample) များကို ရရှိလာပါသည်။

Continuous and Discrete Data

- Sample Data များသည် Continuous သို့မဟုတ် Discrete Data များဖြစ်နိုင်ပါသည်။
- အလေးချိန်၊ အရပ် စသည်တို့သည် Continuous Data များဖြစ်ကြပြီး လူဦးရေ၊ မိသားစုဦးရေတို့သည် Discrete Data များဖြစ်ကြသည်။
- Continuous Data = {0.001, 0.0011, 0.0012, 0.0013, ...}
- Discrete Data = {2, 3, 4, 5, 6, ... }
- ကျွန်တော်တို့ ဒီ Data များကို ဘာလုပ်ကြမည်လဲ။
- ကျွန်တော်တို့သည် အများအားဖြင့် Data များကို အမျိုးအစားခွဲခြင်း (Classification)၊ အလယ်ညွှန်းကိန်းများ ရှာခြင်း (Central Tendency)၊ သွေဖယ်ခြင်း (Variation)နှင့် ပျံ့နှံ့ခြင်း (Distribution)များကို ရှာခြင်း၊ ကြိမ်နှုန်းများ (Frequency)ကို ရှာခြင်း၊ ပုံမမှန်ခြင်း (Outliers) များကို ရှာခြင်း တို့ကို ပြုလုပ်ကြပါသည်။
- ဒီလို Data များ အပေါ် လုပ်ဆောင်သော သင်္ချာနည်းစနစ်များကို Statistics ဟုခေါ်ပါသည်။

Classification

- မြို့တစ်မြို့ရှိ လူဦးရေ၏ အသက်များ ဆိုပါတော့။ ကျွန်တော်တို့ လူ 41 ကို Random Sample လုပ်ထားသည် ဆိုပါတော့။

```
Age = [2, 3, 4, 5, 3, 2, 1, 1, 4, 15, 14, 14, 14, 13, 12, 18, 18, 19,
20, 20, 23, 21, 24, 30, 32, 34, 36, 32, 33, 41, 40, 42, 35, 50, 65, 54,
61, 73, 34, 63, 100]
```

- ကျွန်တော်တို့ အသက်များကို ဒီအတိုင်းထားမည့်အစား Class (Group) များလုပ်ပြီး Classification လုပ်လို့ရပါသည်။

```
Age 1 - 10       : 1, 1, 2, 2, 3, 4, 4, 5
Age 10 - 20      : 12, 13, 14, 14, 14, 15, 18, 18, 19, 20, 20
Age 20 - 30      : 21, 23, 24, 30
Age 30 - 40      : 32, 32, 33, 34, 34, 35, 36, 40
Age 40 - 50      : 41, 42, 50
Age > 50         : 54, 61, 63, 65, 73, 100
```

Central Tendency

- ပြီးရင် Central Tendency ကို ကြည့်ရအောင်ပါ။ Central Tendency ကို Mean (Average), Median, Mode တို့ဖြင့် ဖော်ပြပါသည်။

```
Age = [2, 3, 4, 5, 3, 2, 1, 1, 4, 15, 14, 14, 14, 13, 12, 18, 18, 19,
20, 20, 23, 21, 24, 30, 32, 34, 36, 32, 33, 41, 40, 42, 35, 50, 65, 54,
61, 73, 34, 63, 100]
```

- Sample Data ၏ Average ကို Sample Mean ဟုခေါ်ပြီး ၎င်းကို $\bar{x}$ ဖြင့်ဖော်ပြပါသည်။
- Data များကို Sort by Ascending လုပ်ပြီး ရရှိလာသော အလယ်ကိန်းကို Median ဟုခေါ်ပါသည်။ အလယ်ကိန်းက 2 ခုဆိုပါက ပေါင်းပြီး 2 ဖြင့် စားပါသည်။
- အကြိမ်အရေအတွက် အများဆုံးရှိသော Sample ကို Mode ဟုခေါ်ပါသည်။

Mean = $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ = 28.5

Median = 23

Mode = 14

More Central Tendency

- Central Tendency သည် Similar Samples from Different Population များကို Compare လုပ်ရာမှာ အင်မတန် အသုံးဝင်ပါသည်။
- ဥပမာ၊ သင်္ချာအတန်း 2 တန်းဆိုပါတော့။ ပထမ အတန်းမှာ သင်္ချာ 100 ရသူ 2 ယောက်၊ 80 အထက် 3၊ ကျန်သူက 70 ရသည် ဆိုပါတော့။ ဒုတိယ အတန်းမှာလည်း သင်္ချာ 100 ရသူ 2 ယောက်၊ 80 အထက် 3 ယောက်၊ ကျန်သူက 70 ရသည် ဆိုပါတော့။
- ဒါဆိုရင် အတန်း 2 တန်းလုံး အတူတူ တော်တယ်လို့ ပြောလို့ရမလား။ အမှန်တော့ မပြောနိုင်ပါ။
- တကယ်လို့ ပထမအတန်း ကျောင်းသား၊ ကျောင်းသူ 10 ယောက်ရှိပြီး ဒုတိယအတန်းက ကျောင်းသား၊ ကျောင်းသူက 6 ယောက်ဆိုရင် ဒုတိယအတန်း ပိုတော်တယ်လို့ ပြောလို့ရသည်။
- ပထမအတန်းရဲ Sample Mean က 79 ဖြစ်ပြီး ဒုတိယအတန်းရဲ Sample Mean က 85 ဖြစ်သည်။
- ထို့ကြောင့် Similar Samples from Different Population များကို Compare လုပ်ရာမှာ Central Tendency ကို သုံးပါသည်။
- Sample Mean သည် Sample Set တစ်ခု၏ Center ကို ဖော်ပြသည်။

Variation

- Variation သည် Central Tendency မှ အကွာအဝေး (Distance)ကို တိုင်းတာခြင်း ဖြစ်သည်။

Square Distance = $(y - x)^2$

Square Distance from Center = $(\bar{x} - x_i)^2$

Variance = Average Square Distance from Center = $\frac{1}{n}\sum_{i=1}^{n}(\bar{x} - x_i)^2$

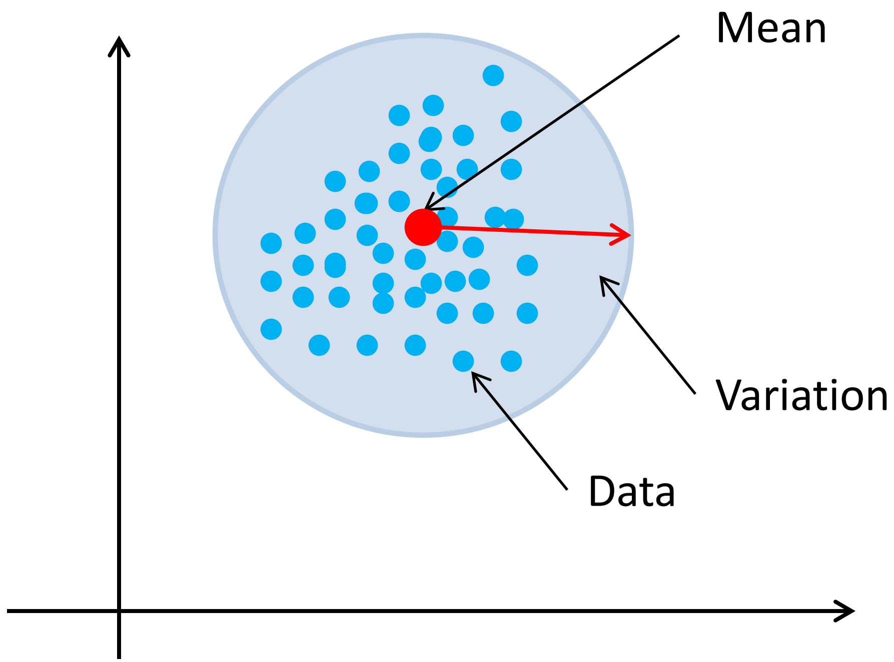
Standard Deviation = Average Distance from Center = $\frac{1}{n}\sum_{i=1}^{n}\sqrt{(\bar{x} - x_i)^2}$

- ဆိုလိုသည်က Sample Mean သည် Sample Set တစ်ခု၏ Center တစ်ခုဆိုပါက Data များသည် Center မှ ဘယ်လောက် အကွာအဝေးတွင် ရှိကြသည်ကို Variation ဖြင့် ဖော်ပြနိုင်သည်။

Mean and Variation

- Sample Mean သည် Sample Set တစ်ခု၏ Center တစ်ခုဆိုပါက Data များသည် Center မှ ဘယ်လောက် အကွာအဝေးတွင် ရှိကြသည်ကို Variation ဖြင့် ဖော်ပြနိုင်သည်။

- Circle တစ်ခုကို Center နှင့် Radius ဖြင့် ဖော်ပြနိုင်သလို Data များကိုလည်း Mean နှင့် Variation ဖြင့် ဖော်ပြလို့ ရသည်။



Frequency

- Frequency (ကြိမ်နှုန်း)သည် Sample Set တစ်ခုတွင် Sample Data များ ဘယ်နှကြိမ် ပါဝင်သည်ကို တိုင်းတာသည်။

```
Age = [2, 3, 4, 5, 3, 2, 1, 1, 4, 15, 14, 14, 14, 13, 12, 18, 18, 19,
20, 20, 23, 21, 24, 30, 32, 34, 36, 32, 33, 41, 40, 42, 35, 50, 65, 54,
61, 73, 34, 63, 100]
```

- Relative Frequency က Frequency ကို Total Number of Sample ဖြင့်စားထားခြင်း ဖြစ်သည်။

$$\text{Relative Frequency} = \frac{f_i}{n}$$

- Cumulative Frequency က Relative Frequency များကို ပေါင်းထားခြင်း ဖြစ်သည်။

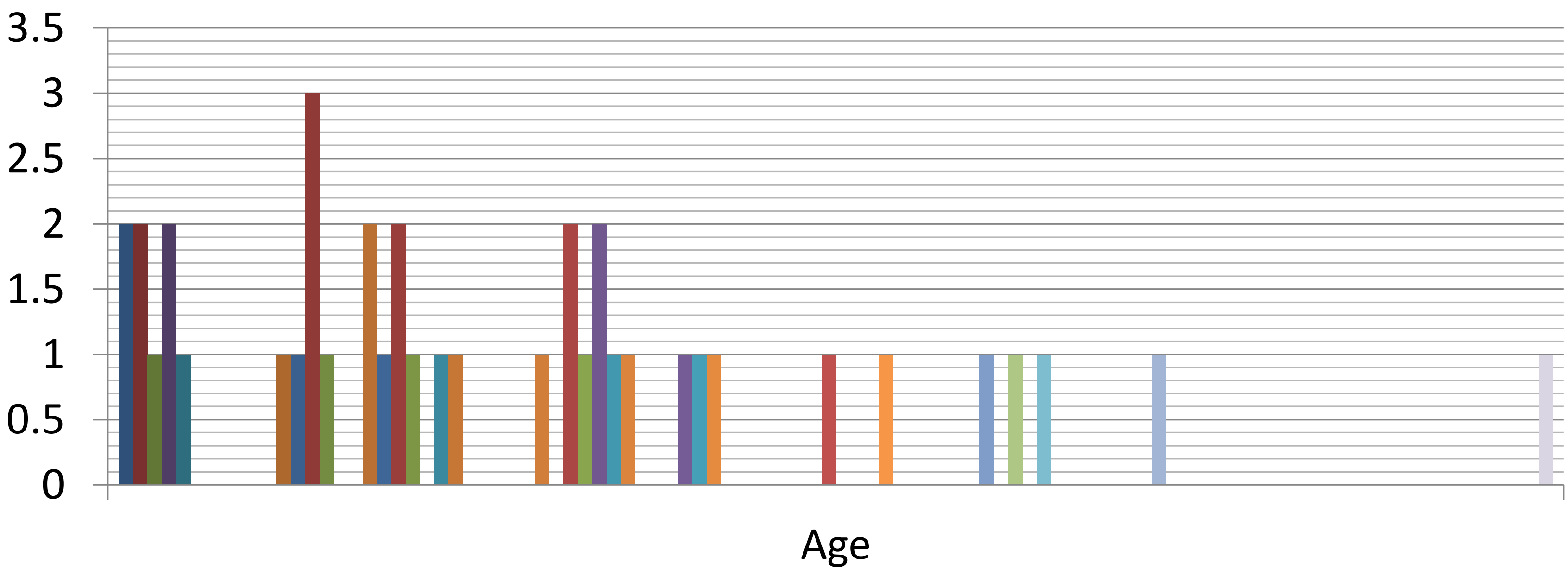
- Sample Mean ကို Frequency မှလည်း ရှာနိုင်ပါသည်။

$$\text{Mean} = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} f_i \times x_i$$

Frequency Distribution

- Frequency နှင့် Age ကို Plot လုပ်ခြင်းဖြင့် Frequency Diagram ကို ရရှိပါသည်။ ထို Diagram သည် Sample Set ၏ Distribution ကို ကိုယ်စားပြုပါသည်။

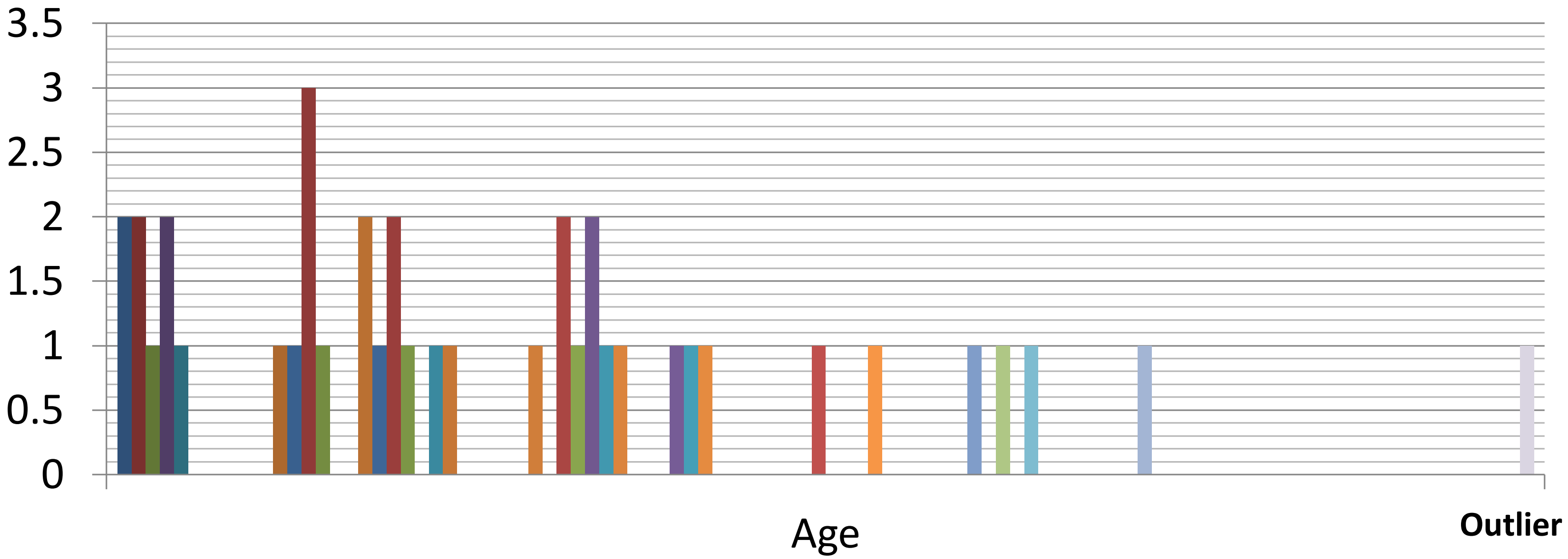
```
Age = [2, 3, 4, 5, 3, 2, 1, 1, 4, 15, 14, 14, 14, 13, 12, 18, 18, 19,
20, 20, 23, 21, 24, 30, 32, 34, 36, 32, 33, 41, 40, 42, 35, 50, 65, 54,
61, 73, 34, 63, 100]
```



Outliers

- Age 100 သည် Outlier တစ်ခု ဖြစ်သည်။ Frequency Diagram မှာ ကြည့်ရင် Age 100 သည် Sample အများစုနှင့် အဝေးမှာ ရှိနေသည်။

```
Age = [2, 3, 4, 5, 3, 2, 1, 1, 4, 15, 14, 14, 14, 13, 12, 18, 18, 19,
20, 20, 23, 21, 24, 30, 32, 34, 36, 32, 33, 41, 40, 42, 35, 50, 65, 54,
61, 73, 34, 63, 100]
```



Statistics Summary

- အခုထိ ပြောပြီး ပြောပြီးတာသည် Basic Statistics ဖြစ်သည်။
- အခုခေတ်ကြီးမှာ Facebook Analytics တို့၊ Google Analytics တို့ ခေတ်စားလာကြပြီး ကိုယ့်ပစ္စည်း သို့မဟုတ် ကိုယ့် Page ကိုကြည့်သူတွေ၊ ဝယ်သူတွေကို Analyze လုပ်လို့ ရလာပါသည်။
- ဥပမာ၊ ကြည့်သူတွေရဲ့ Age သို့မဟုတ် Location ရဲ့ Central Tendency ကို ရှာခြင်းဖြင့် ပုံမှန်ဝင်ကြည့်သူတွေက အသက်ဘယ်လောက်ရှိလဲ၊ ဘယ်က ကြည့်တာများလဲ။ Outliers တွေက ဘယ်သူတွေဖြစ်မလဲ။
- ထို့ကြောင့် Statistics သည် အင်မတန် အသုံးဝင်ပါသည်။

Probability Space

- Probability Space ကို Sample Space (Ω)၊ Random Events (F) နှင့် Probability Measure (P) တို့ဖြင့် ဖော်ပြပါသည်။

- Sample Space (Ω) သည် A Set of Possible (Observable) Outcomes ဖြစ်ပါသည်။ Sample Space သည် Population ရဲ့ Subset တစ်ခု ဖြစ်သည်။

- Random Event သည် Repeat လုပ်နိုင်သော Experiment (Trial) တစ်ခု၏ Outcome ဖြစ်သည်။ ဥပမာ၊ ခေါင်းပန်းလှန်ခြင်းသည် Experiment တစ်ခုဖြစ်ပြီး ထပ်ခါထပ်ခါ Repeat လုပ်နိုင်သည်။ ခေါင်းကျခြင်း၊ ပန်းကျခြင်းတို့သည် Random Event ဖြစ်သည်။ Experiment ၏ Outcome ဖြစ်သော Random Events (F) များသည် Sample Space (Ω) ၏ Subset ဖြစ်သည်။

- Probability Measure က Random Event တစ်ခုစီ၏ Probability ကို ဖော်ပြ (Assign လုပ်)သည်။ Probability Measure တစ်ခုသည် လုံးဝ မဖြစ်နိုင်ခြင်း (0) နှင့် လုံးဝဖြစ်နိုင်ခြင်း (1) တို့ကြား မှာရှိသည်။

- Probability Measure ကို Experimental Estimation သို့မဟုတ် Theoretical Estimation တို့နှင့် Assign လုပ်နိုင်သည်။

Experimental Estimation of Probability

- ခေါင်းပန်းလှန်ရင် တကယ်တော့ ဘာကျမလဲ ဘယ်သူမှ သေချာမသိ။ ခေါင်းကျမည်၊ ပန်းကျမည်၊ ဒေါင်လိုက်ကျမည်၊ ဒင်္ဂါး 2 ခြမ်းလည်းကွဲသွားနိုင်သည်။ ဖြစ်နိုင်သော အရာအားလုံးသည် Population ဖြစ်သည်။ ထိုအထဲမှ စိတ်ဝင်စားသော ဖြစ်နိုင်ချေများသည် Sample Space ဖြစ်သည်။ ခေါင်းကျခြင်းနှင့် ပန်းကျခြင်း။

- Experimental Estimation မှာ Sample Space နှင့် Population တို့ မတူဟု ယူဆထားသည်။

- ထို့ကြောင့် ခေါင်းကျခြင်း၊ ပန်းကျခြင်းသည် ဒီအတိုင်း စဉ်းစားလို့ မရတော့။ ခေါင်းပန်း လှန်ကြည့်ရမည်။ ဒါကို Experiment (Trial) ဟုခေါ်သည်။ Experiment တစ်ခု၏ Outcome သည် Random Event တစ်ခု ဖြစ်သည်။ Random Event သည် Sample Space ၏ Subset တစ်ခု ဖြစ်သည်။

- Random Event တစ်ခုမှာ Probability Measure တစ်ခု ရှိသည်။ ဥပမာ၊ ခေါင်းပန်း 10 ကြိမ်မှာ ခေါင်းကျခြင်း P(H) သည် 0.43234 ဖြစ်သည်။

- Random Event တစ်ခု၏ Probability Measure by Experimental Estimation သည် လက်တွေ့ လုပ်ဆောင်ခြင်းမှ ရရှိလာသည်။

Theoretical Estimation of Probability

- Theoretical Estimation မှာ Sample Space နှင့် Population တို့ တူသည် ဟု ယူဆထားသည်။

- ထို့ကြောင့် ခေါင်းကျခြင်း၊ ပန်းကျခြင်း စသည့် Random Event ၏ Probability Measure ကို Estimate လုပ်ရတာ လွယ်ကူသွားသည်။ ဥပမာ၊ P(H) သည် 0.5 ဖြစ်သည်။

- Random Event တစ်ခု၏ Probability Measure by Theoretical Estimation သည် Theoretical Assumption တစ်ခုသာ ဖြစ်သည်။

- လက်တွေ့မှာ ခေါင်းကျခြင်း၏ Probability P(H) သည် 0.5 အတိအကျ မဟုတ်နိုင်။ ဥပမာ၊ ခေါင်းပန်း 2 ခါလုပ်လို့ ပထမတစ်ခါ ခေါင်းကျပြီး ဒုတိယ ပန်းကျမည်လို့ မပြောနိုင်။

- သို့သော် တွက်ချက်ရလွယ်အောင် Random Event တစ်ခု၏ Probability Measure by Theoretical Estimation ကို ကျွန်တော်တို့ လောလောဆယ် စဉ်းစားပါမည်။

- ထို့ကြောင့် Probability ဟုပြောလျှင် Probability Measure by Theoretical Estimation ကို ဆိုလိုပါသည်။ Theoretical Estimation နှင့် Experimental Estimation တို့ ဆက်စပ်မှုကို နောက်ပိုင်း ဆွေးနွေးပါမည်။

Probability

- ကျွန်တော်တို့ ငယ်ငယ်တုံးက Probability ကို ဒီလို define လုပ်ခဲ့ပါသည်။

$$P = \frac{\text{A set of favorable outcomes}}{\text{A set of possible outcomes}}$$

- ခေါင်းပန်းလှန်ရင် A set of possible outcomes = {H, T}၊ A set of favorable outcomes = {H}။ ဒါကြောင့် P(H) = 1 / 2 = 0.5 လို့ တွက်ခဲ့ကြပါသည်။

- ဒါဆိုရင် မနက်ဖြန် မိုးရွာမလား ဆိုပါတော့။ A set of possible outcomes = {Rain, No Rain}၊ A set of favorable outcomes = {Rain}။ ဒါကြောင့် P(Rain) = 1 / 2 = 0.5 ။ ဒါကြောင့် မနက်မိုးရွာနိုင်ခြေ 50 ရာခိုင်နှုန်း ရှိသည်ဆိုရင် မှန်မလား။

- ဒီအချက် အင်မတန် အရေးကြီးပါသည်။ အမှန်က Probability ကို ရရှိတာသည် Probability Distribution က ရရှိတာ ဖြစ်ပါသည်။ ဒါကြောင့် အထက်ပါ Formula သည် Uniform Probability Distribution အတွက်ပဲ မှန်ပါသည်။

- Uniform Probability Distribution မှာ Sample Space ရှိ Sample များသည် ဖြစ်နိုင်ခြေအတူတူ (Equally Likely) ဖြစ်လို့ ဖြစ်သည်။ အခြား Probability Distribution များ သည် Equally Likely မဖြစ်ပါ။ ဒါကြောင့် မိုးရွာခြင်း၊ မရွာခြင်းသည် Uniform Probability Distribution မဟုတ်ပါ။

- ဒီအချက်က အင်မတန် အရေးကြီးပါသည်။ ဒါကို မသိရင် အလိမ်ခံရနိုင်သည်။ ဒါကြောင့် Probability လို့ ပြောရင် Equally Likely နှင့် Probability Distribution ကို စဉ်းစားဖို့ လိုအပ်သည်။

Probability

- ထို့ကြောင့် Sample Space (A set of possible outcomes) အတွင်းရှိ Sample များသည် ဖြစ်နိုင်ခြေအတူတူ (Equally Likely) ဖြစ်မှသာလျှင်

$$P = \frac{\text{A set of favorable outcomes}}{\text{A set of possible outcomes}}$$

- ဒီအတွက် Condition 2 ခု ရှိရပါသည်။
  - (၁) Sample Space (A set of possible outcomes) ကို သိရပါမည်။
  - (၂) Sample Space အတွင်းရှိ Sample များသည် ဖြစ်နိုင်ခြေအတူတူ (Equally Likely) ဖြစ်ရပါမည်။

> "If we knew Lady Luck better, Las Vegas would still be a road-stop in the desert."
>
> *-Stephen Jay Gould*

Expectation

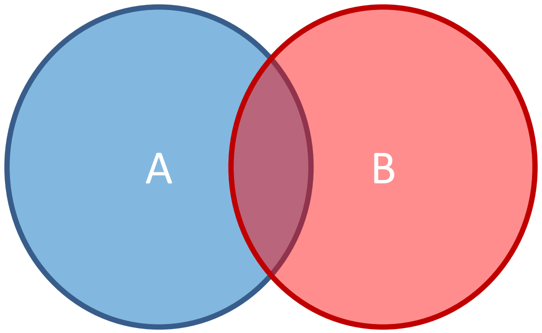
- Probability Measure ကို Experimental Estimation သို့မဟုတ် Theoretical Estimation တို့နှင့် Assign လုပ်နိုင်သည်လို့ ကျွန်တော်တို့ ပြောခဲ့ပါသည်။
- အမှန်တော့ Statistics သည် Empirical (Experimental) Framework (အထူးသဖြင့် Sample Space) ကို အဓိကထားတာ ဖြစ်ပြီး Probability က Theoretical Framework (Sample Space and Population) ကို အဓိက ထားတာ ဖြစ်သည်။
- ထို့ကြောင့် Statistics နှင့် Probability တို့သည် ဆက်စပ်လျက် ရှိသည်။
- Statistics ရှိ Relative Frequency သည် Probability Measure by Experimental Estimation သို့မဟုတ် Empirical Probability ဖြစ်သည်။
- ထိုနည်းတူ Statistics ရှိ Mean (Average) သည် Probability အရ Expected Value ဖြစ်လာသည်။

$$\text{Expectation} = E(x) = \sum_{i=1}^{n} P_i \times x_i$$

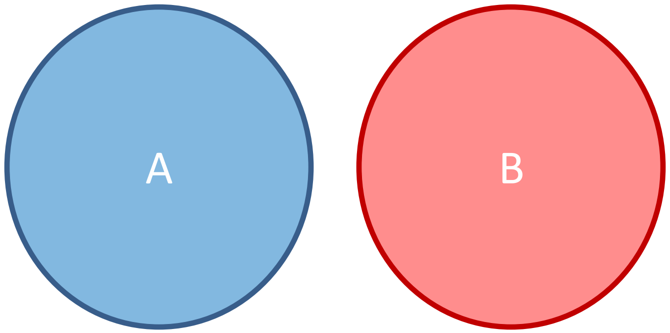
Exclusivity

- Random Event များသည် တစ်ခု ဖြစ်လျှင် တစ်ခု ဖြစ်လို့မရပါက ထို Random Event များကို Exclusive (Mutually Exclusive) ဖြစ်သည်ဟု ခေါ်သည်။

- Random Event များသည် Exclusive မဖြစ်ပါက Inclusive ဖြစ်ကြသည်။



Inclusive

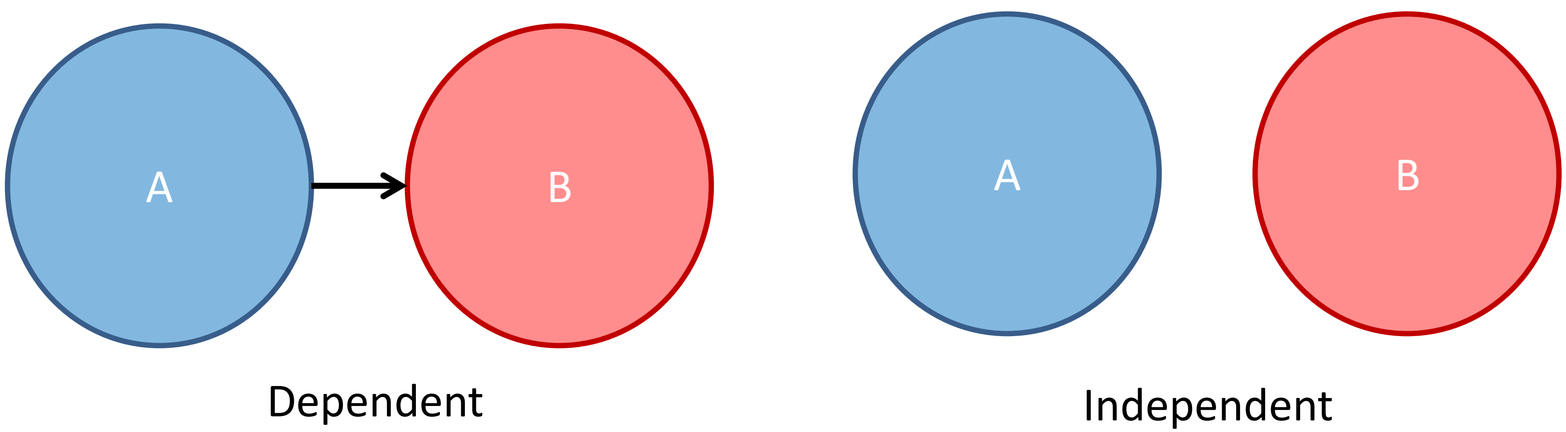


Exclusive

Dependency

- Random Event တစ်ခု၏ Probability Measure သည် အခြား Random Event တစ်ခု အပေါ်မူတည်ပါက Random Event များသည် Dependent ဖြစ်ကြသည်။

- Random Event များသည် Dependent မဖြစ်ပါက Independent ဖြစ်ကြသည်။



Dependency = P(B/A) = Probability of B given A

Operations of Random Events

Operations	**Exclusive**	**Inclusive**
P(A) ∪ P(B)	P(A) + P(B)	P(A) + P(B) – P(AB)
P(A) ∩ P(B)	P(AB) = 0	P(A) . P(B)
P(A’)	1 – P(A)	1 – P(A)
P(A) – P(B)	P(A) – P(B)	P(A) – P(AB)

Operations	**Dependent**	**Independent**
P(A) ∩ P(B)	P(A | B) . P(B) = P(B | A). P(A)	P(A). P(B) = P(AB)
P(A’)	1 – P(A)	1 – P(A)
P(A | B)	P(AB) / P(B)	P(A)
P(B)	P(B | A)P(A) + P(B | A’)P(A’)	P(B)

Bayes’ Theorem

- Random Event တစ်ခု၏ Probability Measure သည် အခြား Random Event တစ်ခု အပေါ်မူတည်ပါက Random Event များသည် Dependent ဖြစ်ကြသည်။

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}$$

- Bayes’ Theorem ကို ဒီက ရပါသည်။

$$P(A) \square\ P(B) = P(A \mid B)P(B) = P(B \mid A)P(A)$$

- Bayes Rule သည် ကျွန်တော်တို့ နောက်ပိုင်းမှာ Statistical Machine Learning အတွက် အင်မတန် အရေးကြီးသော အခြေခံ ဖြစ်ပါသည်။

Discriminative and Generative Model

- Discriminative နှင့် Generative Model သည် Statistical Machine Learning မှာ အသုံးများပါသည်။
- Discriminative Model အရ Statistical Data များမှ ကျွန်တော်တို့က Joint Probability P(X, Y) နှင့် Priori Probability P(X) ကို သိပါသည်။ ဒါကို အခြေပြုပြီး P(Y | X) ကို Bayes Rule ပေါ် အခြေခံပြီး ပြန်တွက်ခြင်း ဖြစ်သည်။

$$P(Y \mid X) = \frac{P(X,Y)}{P(X)}$$

- Generative Model အရ Statistical Data များမှ ကျွန်တော်တို့က Marginal Probability P(Y) နှင့် Posterior Probability P(X |Y) ကို သိပါသည်။ ဒါကို အခြေပြုပြီး Joint Probability P(Y, X) ကို ပြန်တွက်ခြင်း ဖြစ်သည်။

$$P(Y, X) = P(Y)\,P(X \mid Y)$$

Example

- ကျွန်တော်တို့ မီးပွိုင့်မှာ ကားတိုက်ခံရခြင်း အကြောင်းကို စဉ်းစားရအောင်ပါ။ ကျွန်တော်တို့မှာ P(H) = Probability of being hit by a car နှင့် P(L) = Probability of traffic light တို့ ရှိပါသည်။
- P(L) ၏ Sample Space သည် {R, Y, G } ဖြစ်ပြီး သူ့ရဲ Probability ကို ကြိုသိပါသည်။ ဒါကြောင့် P(L) သည် Priori Probability (ကြိုသိထားသော Probability) ဖြစ်သည်။ သို့သော် P(H) ကို ကျွန်တော်တို့ မသိပါ။ ထို့ကြောင့် P(H) သည် Marginal Probability ဖြစ်သည်။ လိုချင်တဲ့ Probability ဖြစ်သည်။

$$P(L = R) = 0.2,\ P(L = Y) = 0.1,\ P(L = G) = 0.7$$

- ကျွန်တော်တို့ မီးနီ၊ မီးဝါ၊ မီးစိမ်းချိန်မှာ ကားတိုက်ခံရခြင်းကို ဖော်ပြသော Conditional Probability P(H | L)ကို လည်း သိပါသည်။

	Red (R)	Yellow (Y)	Green (G)
Not Hit	0.99	0.9	0.2
Hit	0.01	0.1	0.8

Example

- အခု ကျွန်တော်တို့က Priori Probability P(L) နှင့် Conditional Probability P(H | L) တို့ပေါ်မူတည်ပြီး Joint Probability P(H, L) = P(HL) ကိုတွက်ကြည့်ပါမည်။

$$P(H, L) = P(H \mid L)\, P(L)$$

- နောက်ဆုံးကျွန်တော်တို့ Marginal Probability P(H) ကို ရရှိပါသည်။

	Red (R)	**Yellow (Y)**	**Green (G)**	**P(H)**
Not Hit	0.198	0.09	0.14	0.428
Hit	0.002	0.01	0.56	0.572
P(L)	0.2	0.1	0.7	1.0

- Marginal Probability P(H) အပေါ်အခြေခံပြီး Posterior Probability P(L | H) ကို တွက်ကြည့်လို့ ရပါသည်။

$$P(L \mid H) = \frac{P(H \mid L)\, P(L)}{P(H)}$$

Experimental and Theoretical Framework

- ကျွန်တော်တို့ ခေါင်းပန်းလှန်ခြင်း နှင့် မိုးရွာမရွာ ခန့်မှန်းခြင်းမှ P(H), P(T) နှင့် P(Rain), P(No Rain) တို့ မတူညီကြောင်း သိခဲ့ပါသည်။

- အခုချိန်ထိ ကျွန်တော်တို့ Probability ကို Probability အနေဖြင့်သာ နားလည်ထားတာ ဖြစ်သည်။ ဒါပေမယ် Probability များသည် တစ်ခုနှင့် တစ်ခု မတူညီကြောင်းတော့ ထိုက်သင့်သလောက် သိလာပြီ ဖြစ်သည်။

- ထို့ကြောင့် Probability များသည် တစ်ခုနှင့် တစ်ခု မတူညီကြောင်းကို လေ့လာဖို့ အထူးလိုအပ်လာပါသည်။ ဒါကြောင့် Probability များကို လေ့လာဖို့ Theoretical Framework တစ်ခု ရှိဖို့ လိုအပ်လာသည်။

- ထို Theoretical Framework အရ ကျွန်တော်တို့သည် Probability များကို Mathematical Function များအနေဖြင့် လေ့လာကြဖို့ ဖြစ်သည်။ Mathematical Function များ အနေဖြင့် လေ့လာခြင်းသည် Function များ၏ ကွဲပြားပုံကို ပိုပြီး နားလည်နိုင်ပါသည်။ ဥပမာ၊ $y = \sin(x)$ နှင့် $z = x^2 + 3x + 5$ တို့သည် Function များ ဖြစ်ကြသော်လည်း တစ်ခုနှင့် တစ်ခု မတူပါ။

- ဒါကြောင့် Probability Function များကို ဘယ်လို သတ်မှတ်မလဲ။ Probability Function များသည် တစ်ခုနှင့် တစ်ခု ဘာကွာလဲ။ ဒါကို ဆက်လက် ဆွေးနွေးမှာ ဖြစ်သည်။

Random Variable

- အလွယ်ဆုံးပြောရင်တော့ Random Variable များသည် Probability Function တစ်ခု၏ Variable များဖြစ်သည်လို့ ပြောလို့ ရပါသည်။ ဥပမာ၊ x သည် f(x)၏ variable တစ်ခု ဖြစ်သည် ဆိုသလို မျိုးပေါ့။

- သို့သော် ဒါကို ပိုပြီးတော့ Formal ဖြစ်အောင် Define လုပ်ဖို့ လိုအပ်ပါသည်။ Variable လို့ ပြောရင် ကျွန်တော်တို့က Domain ကို ပြောရမှာ ဖြစ်သည်။

- ပုံမှန်အားဖြင့်တော့ Domain ကို Define လုပ်ရတာလွယ်သည်။ ဥပမာ၊ x သည် -400 နှင့် 400 ကြားရှိ Real Number များဖြစ်သည်။

- သို့သော် Probability မှာ ကြတော့ P(Head), P(Rain), P(No Rain), P(being hit by a car)။ ဒါဆိုရင် Probability Function တစ်ခုရဲ့ Domain က ဘာလဲ။

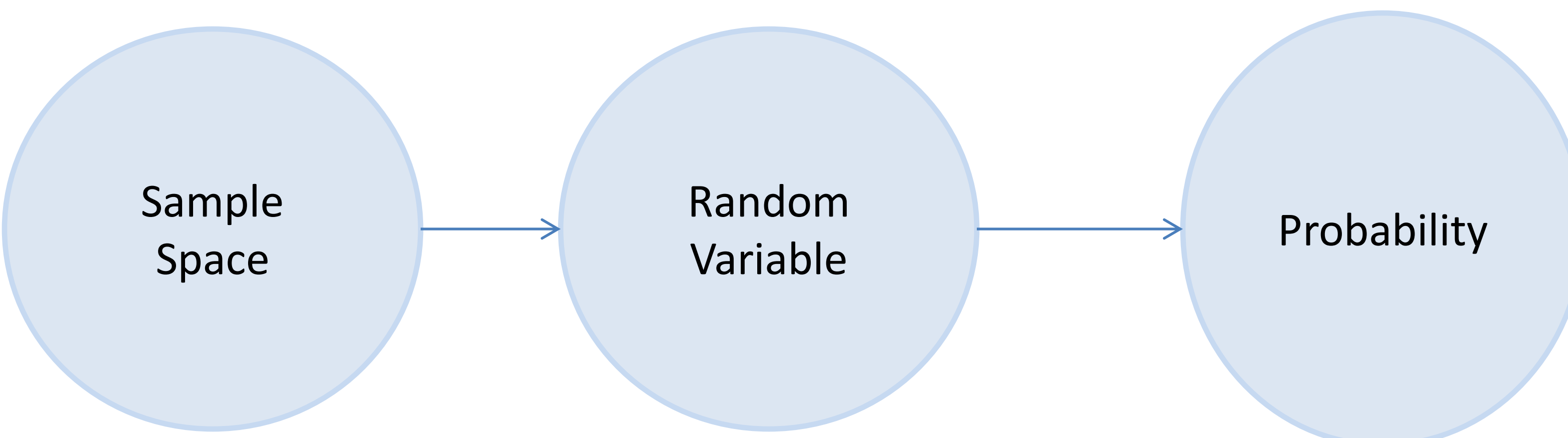
- ဒါကြောင့် Random Variable ကို Set တစ်ခု အနေဖြင့် မယူဆတော့ဘဲ Measurable Function တစ်ခု အဖြစ်ယူဆပါသည်။

> A random variable is understood as a measurable function defined on a probability space that maps from the sample space to the real numbers.

More Random Variable

- တစ်နည်းအားဖြင့် Random Variable သည် Sample Space ကို Real Number (Continuous or Discrete) Map လုပ်ပေးသော Function တစ်ခု ဖြစ်ပြီး Probability Function တစ်ခု၏ Domain လည်းဖြစ်သည်။

- Random Variable ကို “X” ဖြင့် Represent လုပ်ကြပါသည်။ $P(X = 2) = 0.02$ ။



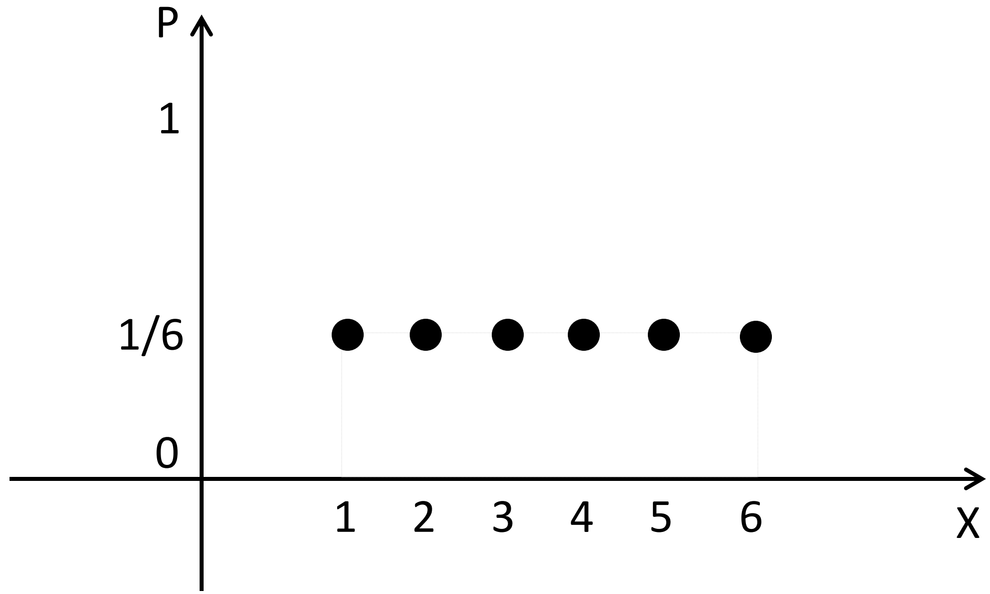
Discrete or Continuous (Real Number)

Probability Distribution

- Probability Distribution ဆိုတာသည် Probability Function တစ်ခု ဖြစ်ပြီး Random Variable များကို Input ယူပြီး Probability ကို Output ထုတ်ပေးသော Function ဖြစ်သည်။

- Probability Distribution များသည် Continuous Function သို့မဟုတ် Discrete Function ဖြစ်နိုင်ပါသည်။

- အန်စာတုံးခေါက်ခြင်း၏ Probability Distribution ကို ဒီလို ဖော်ပြနိုင်သည်။ အန်စာတုံးခေါက်ခြင်း၏ Probability Distribution သည် Uniform Probability Distribution [P(X) = 1/6] ဖြစ်သည်။ Function အနေဖြင့် ပြောရင် Constant Function ဖြစ်သည်လို့ ဆိုလို့ရနိုင်ပါသည်။)



Probability Distributions

- Uniform Probability Distribution သည် အရိုးရှင်းဆုံးသော Probability Distribution ဖြစ်ပြီး Constant Function တစ်ခုနှင့် အလားတူပါသည်။ $y = f(x) = constant; y = P(X) = constant$ ။

- ပိုပြီးရှုပ်ထွေးသော Mathematical Function များရှိသလို ပိုပြီးရှုပ်ထွေးသော Probability Distribution များလည်းရှိပါသည်။

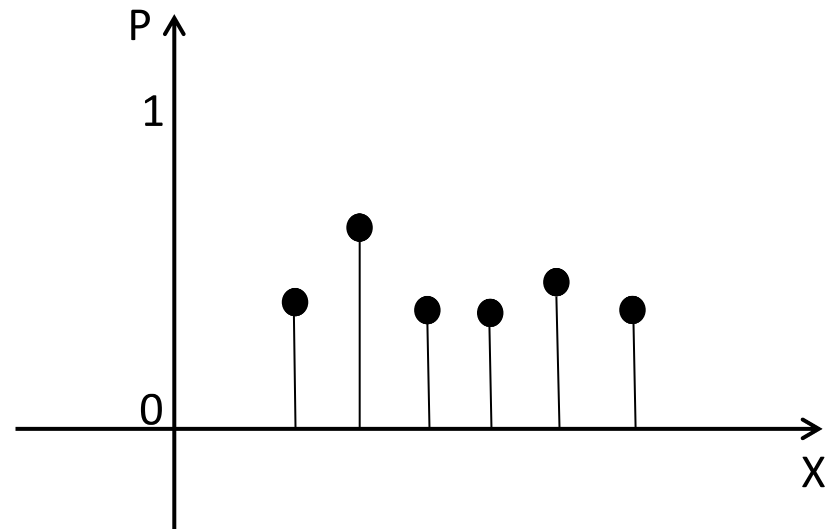
- အသုံးများဆုံးက Distribution များက
  - Uniform Distribution
  - Bernoulli Distribution
  - Normal (Gaussian) Distribution
  - Binomial Distribution
  - Poisson Distribution
  - Exponential Distribution

- ဒါတွေကို ကျွန်တော်တို့ ဆက်ပြီး ဆွေးနွေးပါမည်။

Probability Mass Function

- Probability Distribution ကို ဖော်ပြကြသော်လည် အများအားဖြင့် Cumulative Probability Distribution ကို ပိုသုံးကြပါသည်။ Discrete Cumulative Probability Distribution ကို Probability Mass Function (PMF) ဟုခေါ်သည်။

$$\text{PMF} = \sum_{i=1}^{n} P(x_i)$$



- Probability ကို PDF မှ တွက်လို့ရပါသည်။

$$P(X = 2) = \sum_{i=2}^{3} P(X = x_i)$$

- PMF သည် အမှန်တော့ Area Under Probability Distribution ဖြစ်ပြီး Total Area သည် အမြဲ 1 ဖြစ်သည်။

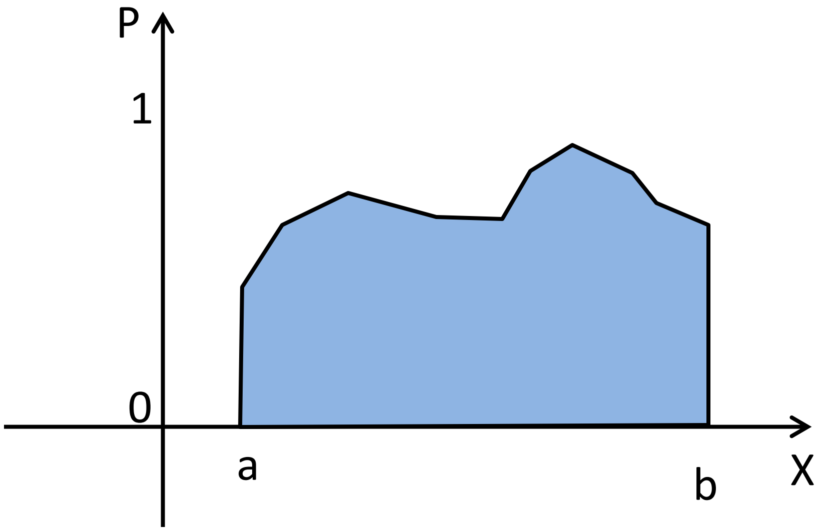
Probability Density Function

- Continuous Cumulative Probability Distribution ကို Probability Density Function (PDF) ဟုခေါ်သည်။

$$\text{PDF} = \int_a^b P(x)dx$$

- Probability ကို PDF မှ တွက်လို့ရပါသည်။

$$P(X = 2) = \int_2^3 P(x)dx$$



Expectation of Random Variable

- Expectation သည် Probability Distribution တစ်ခု၏ Mean (Center) ဖြစ်သည်။
- For Discrete Probability Distribution

$$\mu = E[x] = \sum_{i=1}^{n} P(x_i)x_i$$

- For Continuous Probability Distribution

$$\mu = E[x] = \int_{a}^{b} P(x)xdx$$

Variance of Random Variable

- Variance သည် Probability Distribution တစ်ခု၏ Square Distance from Mean (Center) ဖြစ်သည်။

- For Discrete Probability Distribution

$$\sigma^2 = E[x - \mu]^2 = \sum_{i=1}^{n} P(x_i)(x_i - \mu)^2$$

- For Continuous Probability Distribution

$$\sigma^2 = E[x - \mu]^2 = \int_a^b P(x)(x - \mu)^2 dx$$

Moment of Random Variable

- Moment သည် Probability Distribution တစ်ခု၏ General Form ဖြစ်သည်။
- For Discrete Probability Distribution

$$E[x - c]^n = \sum_{i=1}^{n} P(x_i)\,(x_i - c)^n$$

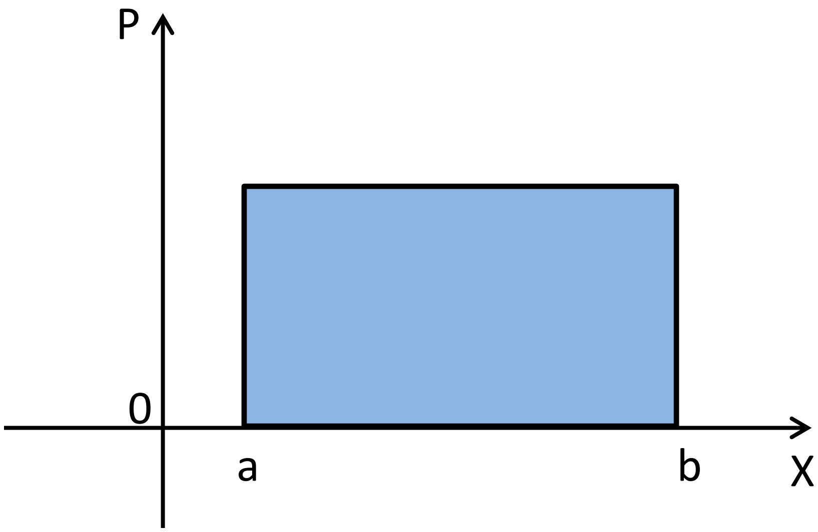
- For Continuous Probability Distribution

$$E[x - c]^n = \int_a^b P(x)(x - c)^n dx$$

- တကယ်လို့ $n = 0$ ဆိုရင် PDF သို့မဟုတ် PMF ကို ရရှိမှာ ဖြစ်သည်။
- တကယ်လို့ $n = 1$, $c = 0$ ဆိုရင် Expectation ကိုရရှိမှာ ဖြစ်သည်။
- တကယ်လို့ $n = 2$, $c = \mu$ ဆိုရင် Variance ကို ရရှိမှာ ဖြစ်သည်။

Uniform Distribution

- ပထမဆုံး Probability Distribution အများသိပြီးသား ဖြစ်သော Uniform Distribution ဖြစ်ပါသည်။
- ကျွန်တော်တို့ Continuous Version အနေဖြင့်သာ ဖော်ပြပါမည်။ (Discrete Version လည်း ရှိပါသည်။)
- P(X) = k = constant ဖြစ်သည်။
- Uniform Distribution ရှိ Random Variable အားလုံးသည် Same Probability ရှိမည် ဖြစ်သည်။



$$E[X] = (b - a) / 2$$

$$\text{PDF} = \int_a^b P(x)\,dx$$

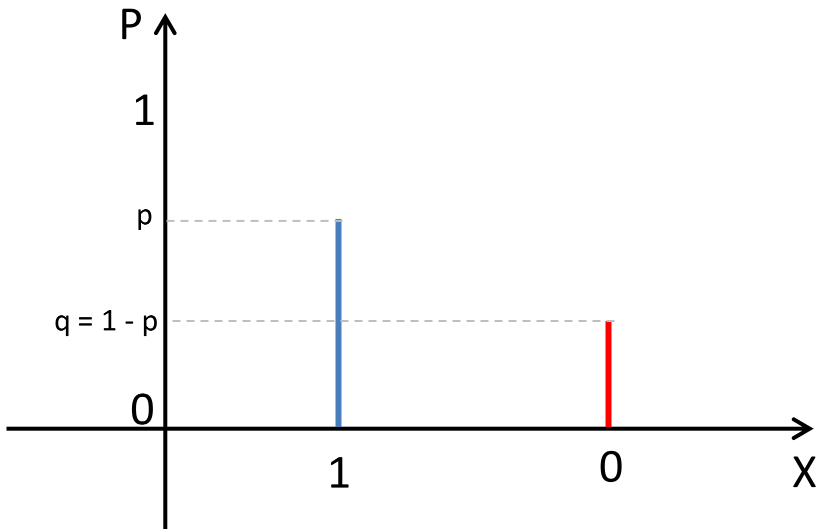
$$VAR[X] = (b - a)^2 / 12$$

Bernoulli Distribution

- Bernoulli Distribution က Discrete Distribution ဖြစ်ပါသည်။ Bernoulli Distribution ကို အထူးသဖြင့် Binary Classification မှာ သုံးပါသည်။

- $P(X = 1) = p$, $P(X=0) = q = 1 - p$ ဖြစ်သည်။ p နှင့် q မတူဘူးလို့ ယူဆထားသည်။ တူသွားရင် Uniform Distribution ဖြစ်သွားမည်။ Bernoulli Distribution ရှိသော Experiment တစ်ခုကို Bernoulli Process ဟု ခေါ်သည်။

- ဥပမာ၊ Email ကို Spam Mail ဟုတ်မဟုတ် Classify လုပ်မည် ဆိုပါတော့။ P(X = spam) = p ဆိုရင် P(X = not spam) = 1 – p ဖြစ်သည်။



$$E[X] = p$$

$$PMF = p^k (1 - p)^{1-k}, \quad k = 0,1$$

$$VAR[X] = pq = p(1 - p)$$

Permutation

- ဂဏန်း 4 လုံးရှိတဲ့ PIN Code တစ်ခု ဆိုပါတော့။ ဂဏန်း တစ်လုံးချင်းသည် 0 မှ 9 (Possible Numbers) အတွင်း ဖြစ်လို့ရသည်။ ဒါဆိုရင် ဂဏန်း 4 လုံး အတွက် All Possible Outcomes သည်

$$P = n^r \quad , n = \text{Possible Number}, r = \text{Length}$$

- ဂဏန်း တစ်လုံးချင်းသည် 0 မှ 9 အတွင်း ဖြစ်လို့ရသည်။ သို့သော် ဂဏန်း 4 လုံးသည် တစ်ခု နှင့် တစ်ခု မတူရဘူး ဆိုပါတော့။ ဒါဆိုရင် ပထမဂဏန်းသည် 10 Possible Numbers၊ ဒုတိယဂဏန်းသည် 9 Possible Numbers…

$$P = n(n - 1)(n - 2)\ldots(n - r + 1) \quad , n = \text{Possible Number}, r = \text{Length}$$

- ဒါကို Factorial ဖြင့် ပြမည်ဆိုရင်

$$^{n}P_{r} = \frac{n!}{(n-r)!} \quad , n = \text{Possible Number}, r = \text{Length}$$

- Permutation သည် Ordered Arrangement များအတွက် သုံးပါသည်။ Permutation (Arrangment) မှာ Order က အရေးကြီးပါသည်။

Combination

- ကျွန်တော်တို့မှာ ဝန်ထမ်းများ { A, B, C, D, E, F, G, H, Z} ရှိသည် ဆိုပါတော့။ ကျွန်တော်တို့က သုံးယောက်ကို အဖွဲ့တစ်ဖွဲ့ (Team or Group) ဖွဲ့ချင်သည်။ ဖြစ်နိုင်တဲ့ All Possible Groups or Teams သည် ဘယ်လို ဖြစ်မလဲ။

- Permutation အရ ကျွန်တော်တို့ (9 × 8 × 7) ရရှိမှာ ဖြစ်ပါသည်။ သို့သော် {A, B, C}, {B, C, A}, {B, A, C}, {C, A, B} တို့သည် Order မတူသော်လည်း တစ်ဖွဲ့ထဲဖြစ်သည်။

- ထို့ကြောင့် Permutation မှ တစ်ဖွဲ့ထဲ ဖြစ်နိုင်သော Arrangement များကို ပြန်ဖယ်ရမည်။ တစ်ဖွဲ့ထဲ ဖြစ်နိုင်သော Permutable Arrangement များသည် r! ဖြစ်သည်။

$$R = r(r-1)(r-2)\ldots(r-r+1) \quad , r = \text{Member}$$

- ထို့ကြောင့် Combination သည် Permutation ကို r! ဖြင့် ပြန်စားထားခြင်းဖြစ်သည်။

$$^{n}C_{r} = \frac{n!}{r!\,(n-r)!} \quad , n = \text{Possible People}, \ r = \text{Member}$$

- Combination သည် Group or Team များအတွက် သုံးပါသည်။ Combination (Grouping) မှာ Order က အရေးမကြီးပါ။

Bernoulli Process

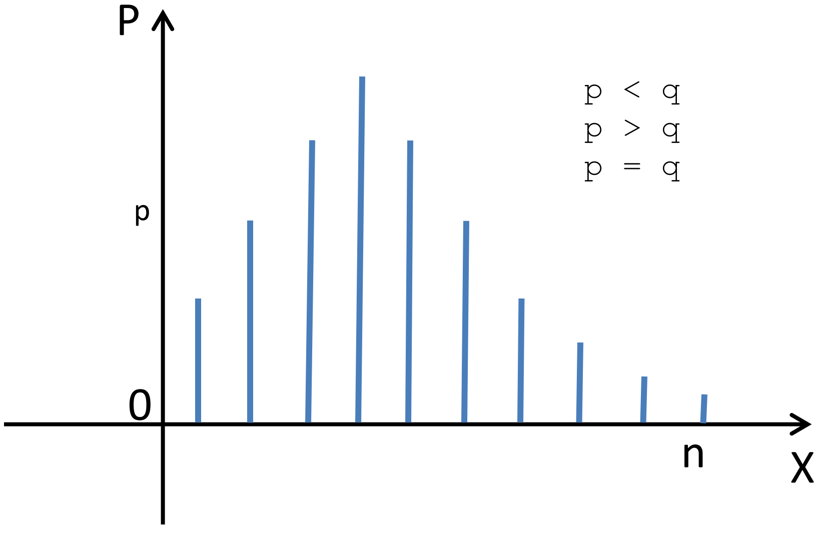
- Experiment တစ်ခုသည် Bernoulli Distribution ရှိသော Bernoulli Process တစ်ခု ဆိုပါတော့။
- ဒီ Experiment ကို N အကြိမ် ထပ်ခါထပ်ခါ လုပ်ဆောင်ကြည့်မည်။ ဒါဆိုရင် N အကြိမ်ထဲမှာမှ ဘယ်နှကြိမ်သည် p [P(X = 1) = p] ရမလဲ။
- ဖြစ်နိုင်သော Random Event များက {p, p, q, q, p, p, ...}, {q, q, p, p, p, ...} စသည်ဖြင့်ပေါ့။ Random Event များသည် p နှင့် q တို့၏ Combination များ ဖြစ်ကြသည်။
- အမှန်တော့ Bernoulli Process ကို N အကြိမ် ထပ်ခါထပ်ခါ လုပ်ဆောင်ကြည့်လို့ ရရှိလာသော p [P(X = 1) = p] ကို Binomial Distribution ဖြင့် ဖော်ပြနိုင်သည်။

Binomial Distribution

- Binomial Distribution က Discrete Distribution ဖြစ်ပါသည်။

- Binomial Distribution ရှိ Random Event တစ်ခုခြင်းစီသည် Bernoulli Process ဖြစ်သည်။ n က ဘယ်နှကြိမ် Repeat လုပ်မလဲ ဖြစ်သည်။

- တကယ်လို့ p: [P(X) = 1] နှင့် q = 1 –p: [P(X) = 0] တို့ တူသွားလျှင် Distribution သည် Symmetric ဖြစ်သည်။ p < q ဆိုလျှင် ဘယ်ဖက် (Skew Left) ကပ်ပြီး p > q ဆိုလျှင် ညာဖက် (Skew Right) ကပ်မည်။



$$E[X] = np$$

$$PMF = {}^{n}C_{r}\, p^{r} (1 - p)^{n-r}$$

$$VAR[X] = npq = np(1 - p)$$

Normal Distribution

- သဲများကို အပေါ်ကနေ သွန်ချလိုက်ပါ။ သဲပုံလေး တစ်ခုရပါမည်။ အလယ်မှာ အများဆုံး ဖြစ်ပြီး ဘေးဖက်များမှာ တစ်ဖြည်းဖြည်းနည်းသွားပါမည်။

- အတန်းထဲရှိ ကျောင်းသားများ၏ အရပ်ကို တိုင်းကြည့်ပါ။ အလယ် (Average) နားမှာ အများဆုံးဖြစ်ပြီး အနိမ့်ဆုံးနှင့် အမြင့်ဆုံးတို့သည် တဖြည်းဖြည်း နည်းသွားမည်။

- နိုင်ငံတစ်နိုင်ငံရှိ အသက်၊ အရပ်၊ လူများ၏ ဝင်ငွေ၊ IQ စသည့် တော်တော်များများသည် သဲပုံ ပုံစံ ဖြစ်နေသည်ကို တွေ့ရပါမည်။

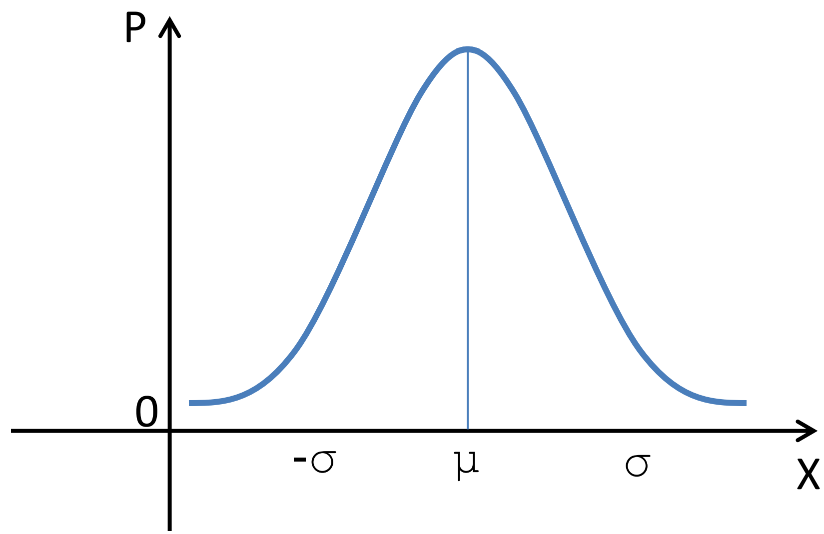
- ဒါသည် ပုံမှန်ဖြစ်လေ့ရှိသော အရာ တစ်ခု ဖြစ်သည်။ အလယ်ရဲ့ Probability သည် ဖြစ်နိုင်ချေ အများဆုံး ဖြစ်သည်။

- ဒါကြောင့်မို့ ဒီလို Probability Distribution ကို Normal Distribution ဟုခေါ်သည်။

- ဒီ Distribution ကို Gauss ကတွေ့ရှိခဲ့တဲ့ အတွက်ကြောင့် Gaussian Distribution လို့လည်းခေါ်ပါသည်။

Normal Distribution

- ကျွန်တော်တို့ Normal Distribution ကို Continuous Version အနေဖြင့်သာ ဖော်ပြပါမည်။ (Discrete Version လည်း ရှိပါသည်။)
- Normal Distribution ၏ Expectation သည် μ (Mean) ဖြစ်သည်။
- Normal Distribution အပါအဝင် အခြား Distribution များသည် Theoretical Sample Space (Population) ကို ပြောတာ ဖြစ်သည်။
- ထို့ကြောင့် Population's Mean μ သည် Sample's Mean $\bar{x}$ နှင့် မတူပါ။



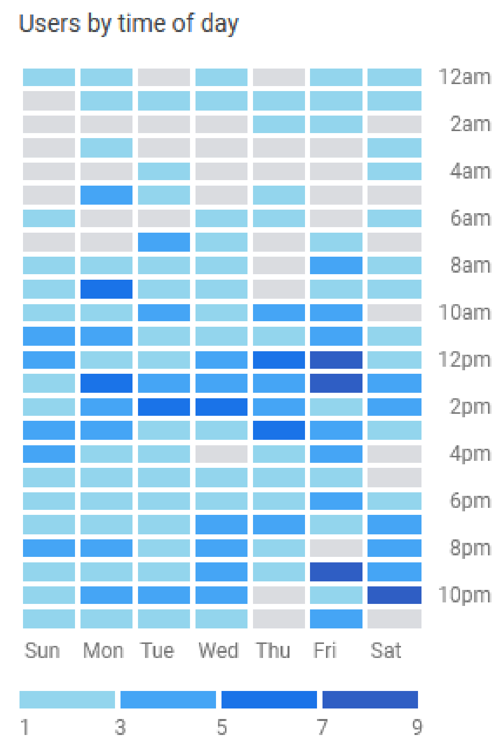
$$E[X] = \mu$$

$$PDF = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$$VAR[X] = \sigma^2$$

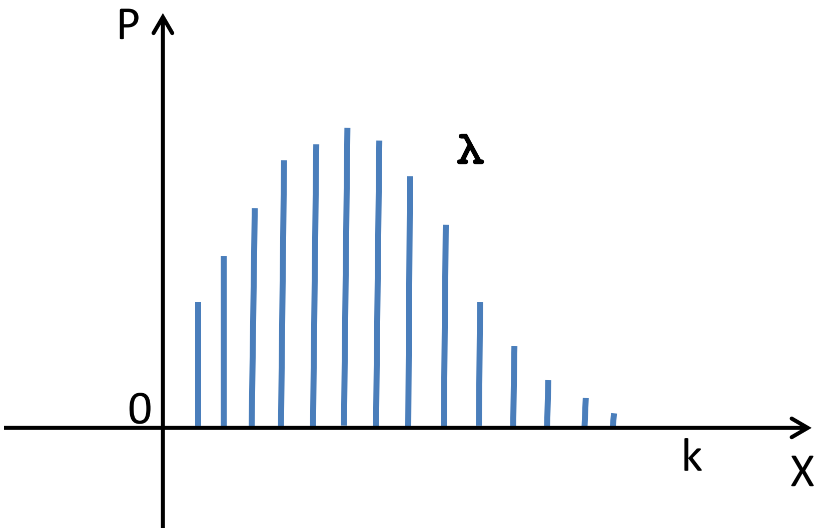
Poisson Distribution

- ခင်ဗျားကို ဆိုင်တစ်ဆိုင် ဖွင့်ထားတယ်ဆိုပါတော့။ Online ပဲ ဖြစ်ဖြစ်၊ Offline ပဲ ဖြစ်ဖြစ်ပေါ့။
- ခင်ဗျားဆိုင်ကို တစ်နေ့ တစ်နေ့ အကြမ်းဖျဉ်း ဝယ်သူ အယောက် 20 လောက်လာ ဝယ်တယ် ဆိုပါတော့။
- ဒါဆိုရင် အခုချိန်မှ စပြီး နောက် 3 နာရီ အတွင်း ဝယ်သူ ဘယ်လောက် လာဝယ်နိုင်မလဲ။
- ဒါရဲ့ အဖြေကို သိချင်ရင်တော့ ကျွန်တော်တို့ Poisson Distribution ကို သုံးရမှာ ဖြစ်ပါသည်။
- အမှန်တော့ Poisson Distribution ကို ဈေးရောင်းမှ မဟုတ်ပါဘူး။ ဘဏ်တွေရဲ့ Teller Counter တွေ၊ ဆေးရုံဆေးခန်းရဲ့ Booking တွေ (Queue လုပ်ရတဲ့ နေရာတိုင်းမှာ) စသည့်နေရာတွေမှာ သုံးနိုင်ပါသည်။
- ကျွန်တော်တို့က Arrival Rate ကိုသာ သိရင် သို့မဟုတ် ခန့်မှန်းနိုင်ရင် ဘယ်အချိန်မှာ Customers ဘယ်လောက်ရှိမယ်ဆိုတာရဲ့ Probability ကို Poisson Distribution နဲ့ ပြောနိုင်ပါသည်။



Poisson Distribution

- Poisson Distribution က Discrete Distribution ဖြစ်ပါသည်။
- λ က Arrival Rate ဖြစ်ပြီးတော့ k က Number of Arrival in a Timer Interval ဖြစ်သည်။
- P(λ = 5, k = 2) ဆိုပါတော့။ Arrival Rate က 5 per day ဆိုပါတော့၊ ဒါဆိုရင် Customer 2 ယောက် 1 နေ့အတွင်းလာနိုင်သော Probability သည် 0.084 ဖြစ်သည်။



E[X] = λ

$$\text{PMF} = \frac{\lambda^k e^{-\lambda}}{k!}$$

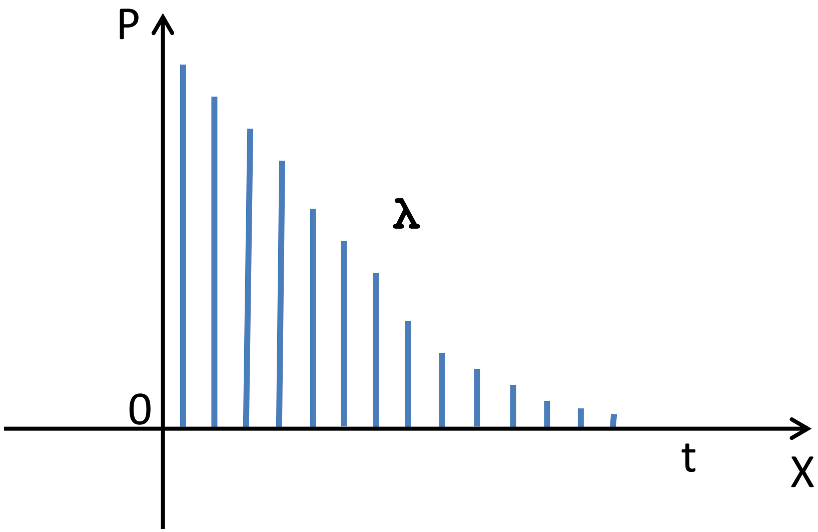
VAR[X] = λ

Exponential Distribution

- ကျွန်တော်တို့ ပြောနေကြရှိပါသည်။ ဆိုင်က အရမ်းရောင်းရတာပဲ၊ လူတွေကို ကြိတ်ကြိတ်တိုးနေတာပဲ။
- အမှန်အားဖြင့် ကြိတ်ကြိတ်တိုးနေရတာ မကောင်းပါ။ Customer များက အမြဲတန်း ကြိတ်ကြိတ်တိုးနေရရင် ကြာရင် တခြားဆိုင်ပြောင်းသွားမှာပါ။ ကြိတ်ကြိတ်တိုးနေရတဲ့ အတွက် Time နှင့် Effort waste ဖြစ်ပါသည်။
- ကျွန်တော်တို့ ဆီမှာက ကြိတ်ကြိတ်တိုးရတာက ထုံးစံလိုပါပဲ။ Passport ရုံးမှာလဲ ကြိတ်ကြိတ်တိုးရသည်။ ဘဏ်မှာလည်း ကြိတ်ကြိတ်တိုးရသည်။ ဘက်စ်ကားလည်း ကြိတ်ကြိတ်တိုးရသည်။ ကြိတ်ကြိတ်တိုးရတာက ရောင်းကောင်းတာကို ပြတာမဟုတ်၊ Service မကောင်းတာကို ပြတာဖြစ်သည်။
- Poisson Distribution က Arrival Rate ကို ပြောပါသည်။ Customer တွေ ရောက်ရှိနှုန်းဖြစ်သည်။ ဒီ Customer များကို Service ပေးရမှာ ဖြစ်သည်။ ဒါကို Service Rate ဟု ခေါ်သည်။ Customer ဘယ်နှယောက်ကို အချိန်တစ်ခုအတွင်း Service ပေးနိုင်လဲ။ Service Rate သည် Arrival Rate ထက်နည်းလို့ ကတော့ ကြိတ်ကြိတ်တိုးပါပြီ။
- ထို့ကြောင့် Arrival Rate ကိုသာ increase လုပ်ဖို့သာ မဟုတ် Service Rate ပါ increase လုပ်ဖို့ လိုအပ်သည်။ ဘဏ်မှာ လူတွေကြိတ်ကြိတ်တိုးရင် Counter အသစ်တိုးခြင်း၊ Branch အသစ်ဖွင့်ခြင်းတို့ လုပ်ရမည်။ ဒါကြီးကို ငါတို့ကတော့ ရောင်းကောင်းလိုက်တာ ဆိုပြီး ဂုဏ်ယူနေဖို့ မဟုတ်။
- Service Rate ကိုသာ သိရင် ကျွန်တော်တို့ Exponential Distribution ကိုသုံးပြီး အချိန်တစ်ခုအတွင်းမှာ Customer ဘယ်နှစ်ယောက်ကို Service ပေးနိုင်ခြင်းရဲ့ Probability ကို တွက်လို့ ရပါသည်။

Exponential Distribution

- Poisson Distribution နှင့် ကိုက်ညီအောင် ကျွန်တော်တို့ Exponential Distribution ကို Discrete Version အနေဖြင့်သာ ဖော်ပြပါမည်။ (Continuous Version လည်း ရှိပါသည်။)
- λ က Service Rate ဖြစ်ပြီးတော့ t က Service Time ဖြစ်သည်။
- P(λ = 5, t = 0.5) ဆိုပါတော့။ Service Rate က 5 Customer per day ဆိုပါတော့၊ ဒါဆိုရင် နေ့တဝက် အတွင်း စောင့်ရမည့် Probability သည် 0.41 ဖြစ်သည်။ ဒါဆိုရင် နေ့တဝက် အတွင်း Service ပေးပြီးသည့် Probability သည် (1 – 0.41) = 0.59 ဖြစ်သည်။



E[X] = $1/\lambda$

PMF = $\lambda e^{-\lambda t}$, t >= 0

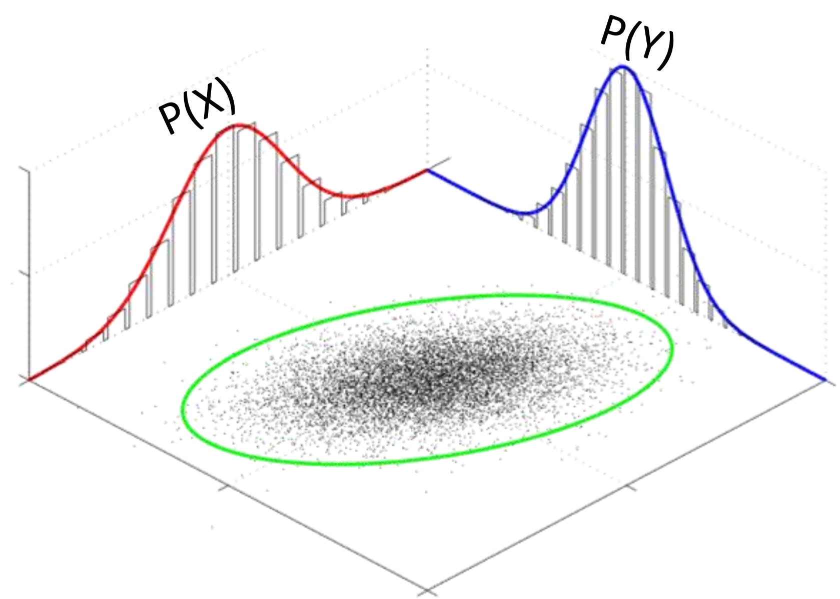
VAR[X] = $1/\lambda^2$

Multivariate Random Variable

- Probability Distribution သည် Function တစ်ခု ဖြစ်သည့် အားလျော်စွာ Random Variable များသည် တစ်ခု ထက်ပို ပါဝင်နိုင်ပါသည်။

$$P(X_1, X_2, X_3, \ldots, X_n)$$

- P(X, Y)၏ Normal Distribution သည် အောက်ပါအတိုင်း ဖြစ်မည်။

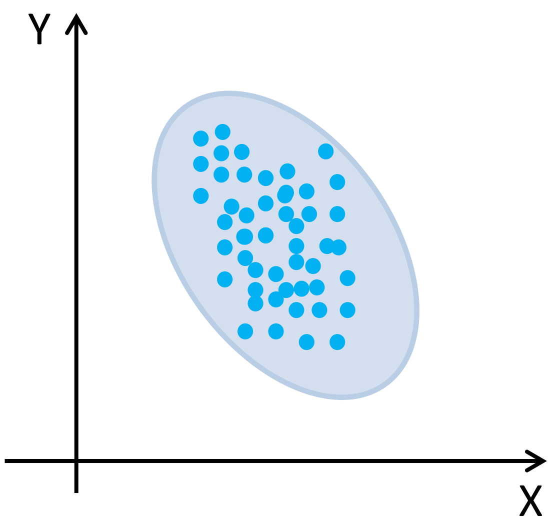


Covariance

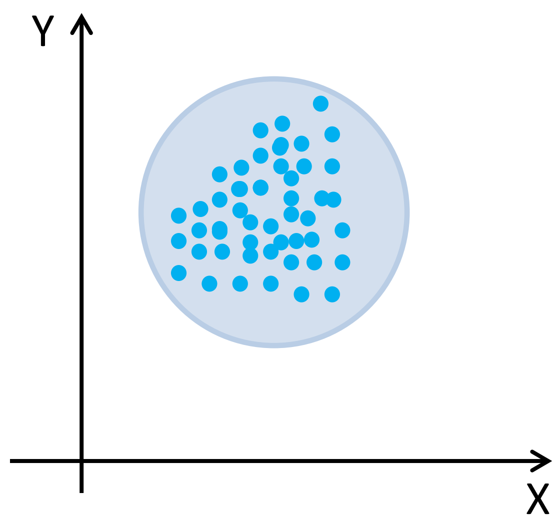
- ကျွန်တော်တို့ Variance ကို Center (Mean) မှ အကွာအဝေးလို့ ပြောခဲ့ပါသည်။ Covariance သည် Joint Variability (Joint Distance) ကို ဖော်ပြဖို့ဖြစ်သည်။

$$COV(X, Y) = E[(X - E[X])(Y - E[Y])]$$

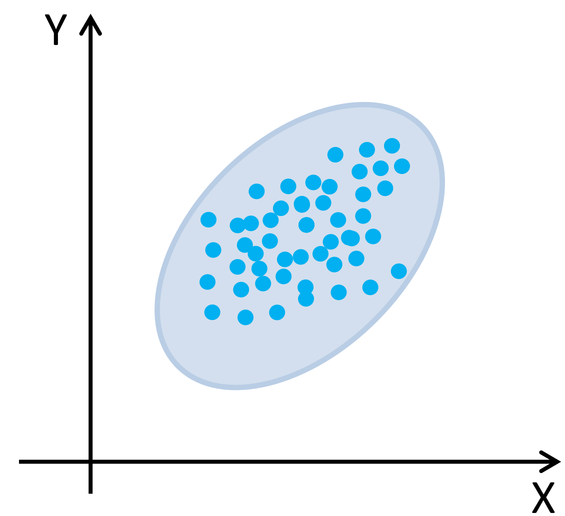
$$COV(X,X) = E[X - E[x]]^2 = VAR(X)$$



COV(X, Y) < 0



COV(X, Y) = 0



COV(X, Y) > 0

Induction and Deduction

- လှပခြင်းသည် ယေဘုယျတရား ဖြစ်ပါသည်။ ကောင်းမွန်ပြည့်စုံခြင်း၊ အပြစ်ကင်းခြင်း၊ ညီညွတ်မျှတခြင်း ဆိုသော ယေဘုယျ အဓိပ္ပါယ်ကို ဆောင်ပါသည်။ လှပခြင်းကို သက်ရှိသတ္တဝါ၊ အရာဝတ္ထု၊ အဖြစ်အပျက်၊ စိတ်သဘောထား စသည့်တို့အား ဖော်ပြကြသည်။ လှပသော ငှက်ကလေးသည်၊ လှပသော တောအုပ်နားရှိ လှပသော အိမ်ကလေးထဲမှ ထွက်လာသည့် လှပသော မိန်းကလေး နားမှာ လှပစွာ ပျံဝဲပြီး တေးသီဆိုနေသည်။ သို့သော် စဉ်းစားဖို့က ငှက်ကလေး၏ လှပခြင်းက မိန်းကလေး၏ လှပခြင်းနှင့် မတူ။ ဒါဆိုရင် ဘာကွဲတာတုံး။

- အသီးတစ်ခုသည် အပင်တစ်ပင်၏ ပန်းပွင့်မှ ဖြစ်လာသော သီးနှံတစ်ခု ဖြစ်သည်။ ပန်းသီး၊ ခရမ်းချဉ်သီးတို့သည် အသီးများ ဖြစ်သည်။ အသီးများကို ခရမ်းချဉ်သီးယို စသည်ဖြင့် ယိုထိုးကြသည်။ အသီးများကို ခရမ်းချဉ်သီးဖျော်ရည် စသည်ဖြင့် ဖျော်ရည်လုပ်ကြသည်။ ထို့ကြောင့် ခရမ်းချဉ်သီးသည် သီးနှံ (Fruit) ဖြစ်သည်။ ဟင်းသီးဟင်းရွက် (Vegetables) မဟုတ်နိုင်။ ဒါဆိုရင် Is Tomato Fruit or Vegetable?

- ထို့ကြောင့် ယေဘုယျ သဘောကို ဖော်ပြသော General Cases များ ရှိသလို တစ်ခုချင်းစီကို ဖော်ပြသော Specific Cases များလည်းရှိပါသည်။ Population သည် General Case ဖြစ်ပြီး Sample Space သည် A Collection of Specific Case ဖြစ်ပါသည်။

- Probability မှလာသော Probability Distribution များသည် General Case ကို ဖော်ပြတာဖြစ်ပြီး Statistics မှလာသော Frequency Distribution များသည် Specific Case ကို ဖော်ပြတာ ဖြစ်သည်။

- ဒါဆိုရင် General မှ Specific၊ Specific မှ General ကို ဘယ်လို ဆက်စပ်ကြပါလဲ။

Induction and Deduction

- Mathematical Reasoning သည် General Case နှင့် Specific Case တို့ ဆက်စပ်ခြင်းကို လေ့လာသော ပညာရပ်ဖြစ်သည်။
- General Case မှ Specific Case ကို စဉ်းစားခြင်းကို Deduction (ဆွဲထုတ်ဆင်ခြင်ခြင်း) လို့ ခေါ်ပြီး Specific Case မှ General Case ကို စဉ်းစားခြင်းကို Induction (ခြုံငုံဆင်ခြင်ခြင်း) လို့ခေါ်သည်။
- အမှန်တော့ သင်ယူခြင်း (Learning) သည် အဆင့် 3 ဆင့် ပါဝင်ပါသည်။ ပထမ ကျွန်တော်တို့ Specific Case (Example) များကို လေ့လာပါသည်။ ဒုတိယ ထို Specific Case များကို General Case များ အဖြစ် Induction (ခြုံငုံဆင်ခြင်ခြင်း) လုပ်ပါသည်။ တတိယ အခြား Specific Case များကို တွေ့သည့်အခါ သင်ယူထားပြီးသော General Case များမှ Deduction (ဆွဲထုတ်ဆင်ခြင်ခြင်း) ပြီး ပြဿနာများကို ရှင်းပါသည်။
- ထို့အတူ Machine Learning လည်းအတူတူ ဖြစ်သည်။ ပထမ ကျွန်တော်တို့ Specific Case (Training Data) များကို Machine များကို သင်ပါသည်။ ဒုတိယ Machine များကို General Case များ အတွက် Model ကို Learn လုပ်စေပါသည်။ တတိယ လိုအပ်သော အခြား Specific Case များကို General Model ဖြင့် ပြန်ဖြေရှင်းပါသည်။
- ထို့ကြောင့် ကျွန်တော်တို့ Specific To General၊ ပြီးရင် General To Specific ကိုဖြေရှင်းရမှာ ဖြစ်သည်။
- Specific To General မှာ ကျွန်တော်တို့က Specific Case ကိုသိသည်၊ General Case ကို ရှာချင်တာဖြစ်သည်။ General To Specific မှာက General Case ကို သိသည်၊ Specific Case ကိုရှာချင်တာ ဖြစ်သည်။

Specific to General Problems

- Specific to General Problems များသည် Deduction ဆိုင်ရာ ပြဿနာများ ဖြစ်ကြသည်။

- ကျွန်တော်တို့ အတန်းထဲရှိ ကျောင်းသာများ၏ အရပ်သည် 5 ပေ 6 လက်မ ပတ်ဝန်းကျင် ဖြစ်သည်။ ဒါဆိုရင် မြန်မာပြည်ရှိ ကျောင်းသားများ၏ အရပ်သည် ဘယ်လောက် ရှိမလဲ။

- Specific to General Problems များကို ဖြေရှင်းနည်း မျိုးစုံ ရှိပါသည်။ ထိုအထဲမှာMaximum Likelihood နှင့် Regression Analysis များကို ပြောပါမည်။

Maximum Likelihood

- ကျွန်တော်တို့ Sampling လုပ်ခြင်းမှ Sample Data များကို ရရှိပါသည်။ ထို Sample Data များ၏ General Probability Distribution သည် ဘာဖြစ်နိုင်မလဲ။

- Probability Distribution တစ်ခုကို Expectation (μ) နှင့် Variance (σ^2) တို့ဖြင့် ဖော်ပြနိုင်ပါသည်။

- Maximum Likelihood မှာ ကျွန်တော်တို့က Sample Data များ၏ Mean ($\bar{x}$) နှင့် Variance (s^2) တို့က ဘယ်လို Probability Distribution Expectation (μ) နှင့် Variance (σ^2) တို့နှင့် အများဆုံး ကိုက်ညီမလဲကို တွက်ချက်တာ ဖြစ်သည်။

- Bayes' Theorem အရ, Likelihood Function သည်

$$L(\theta) = P(X \mid \theta) \quad , X = [x_1, x_2, x_3, \dots, x_n] = \text{constants}, \ \theta = [\mu, \sigma^2]$$

$$\text{Maximum Likelihood :} \quad \frac{dL}{d\theta} \ln\left(L(\theta)\right) = 0$$

Regression Analysis

- Regression Analysis ကို ကျွန်တော်တို့က Linear Regression နှင့် Non Linear Regression ဆိုပြီး ခွဲနိုင်ပါသည်။

- Linear Regression က Advanced Calculus ရဲ့ Supervised Learning မှာ ပြောပြီးသားဖြစ်လို့ ထပ်မပြောတော့ပါ။ အဲဒါလေး ပြန်ဖတ်ကြည့်ပါ။

- ကျွန်တော်တို့က ဒီမှာ Non Linear Regression တစ်ခု ဖြစ်တဲ့ Logistics Regression အကြောင်းကို ပြောပါမည်။

Logistics Regression

- Logistics Regression ကို Binary Classification လုပ်ရာမှာ သုံးပါသည်။ Logistics Regression ကို Binary Regression လို့ ယူဆလို့ ရပါသည်။

- Logistics Function သည် Sigmoid Function ဖြစ်သည်။

$$\sigma(t) = \frac{e^t}{e^t + 1} = \frac{1}{1 + e^t}$$

- ကျွန်တော်တို့မှာ Study Hour (H) နှင့် Exam Pass (P) တို့ရဲ့ Sample Data ရှိသည် ဆိုပါတော့။

H	0.50	0.75	1.00	1.25	1.50	1.75	1.75	2.00	2.25	2.50	2.75	3.00	3.25	3.50	4.00	4.25	4.50	4.75	5.00	5.50
P	0	0	0	0	0	0	1	0	1	0	1	0	1	0	1	1	1	1	1	1

$$P(X = 1 \mid H) = \frac{1}{1 + e^{-(1.5046H - 4.077)}}$$

- Sample များမှ P(X = 1 | H) Logistics Model ကိုရပါသည်။ ထို Logistics Model ကို သုံးပြီး H = 1.8 ကို တွက်ကြည့်လျှင် P(X = 1 | H = 1.8) = 0.2 ဖြစ်သည်။

General to Specific Problems

- General to Specific Problems များသည် Induction ဆိုင်ရာ ပြဿနာများ ဖြစ်ကြသည်။ General to Specific Problems များတွင် ကျွန်တော်တို့သည် General Case များကို သိသည်ဟု ယူဆထားသည်။

- General Case ဆိုသည်မှာ အများအားဖြင့် Theory များဖြစ်ကြသည်။ ထို Theory များကို သုံးပြီး လက်တွေ့ ဘယ်လို ပြန်ရှင်းမလဲ ဆိုတဲ့ ပြဿနာများ ဖြစ်သည်။

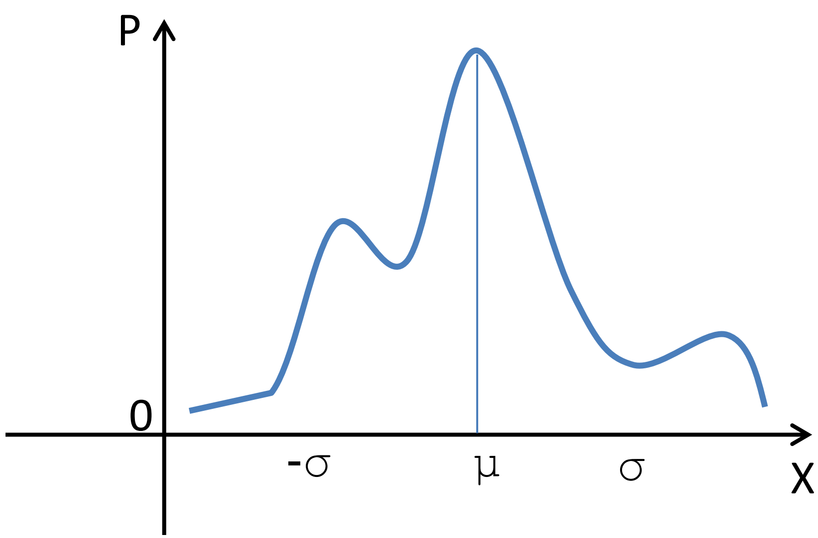
- General to Specific Problems များကို ဖြေရှင်းနည်း မျိုးစုံ ရှိပါသည်။ ထိုအထဲမှာ Hypothesis Testing များကို ပြောပါမည်။

Central Limit Theorem

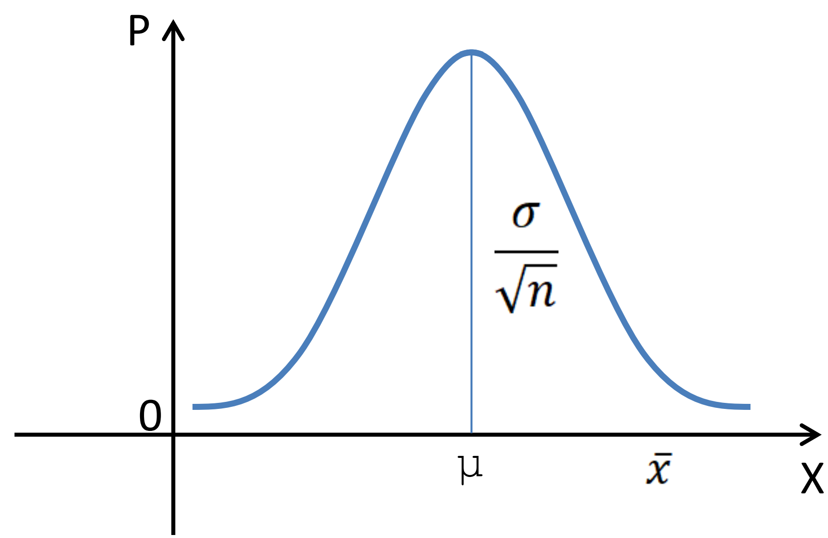
- ကျွန်တော်တို့သည် Expectation (μ) နှင့် Variance (σ^2) ရှိသော Population တစ်ခုမှ Sample များကို ရွေးထုတ်လိုက်သည် ဆိုပါတော့။ ကျွန်တော်တို့ မူလ Population ရဲ့ Probability Distribution ကို မသိပါ။ သို့သော် Randomly ရွေးထုတ်လိုက်သော N Sample များ၏ Mean($\bar{x}$)ရဲ့ Probability Distribution သည် အမြဲ Normal Distribution ဖြစ်သည်။

- ပန်းသီးခြင်းတစ်ခုမှာ ပန်းသီးများ ရှိသည်။ ထိုပန်းသီးများရဲ့ Weight ကို မသိပါ။ ထို့ကြောင့် N ပန်းသီးများကို ရွေးပြီး ချိုတဲ့အသီးများရဲ့ Mean (ပျမ်းမျှ Weight)ကိုရှာပါသည်။ ထို့နောက် နောက်ထပ် N ပန်းသီးများကို ရွေးပြီး ချိုတဲ့အသီးများကို Mean ကိုရှာပြန်ပါသည်။ ထို Sample များ၏ ချိုတဲ့အသီးများ၏ Mean သည် အမြဲ Normal Distribution ဖြစ်သည်။

Original Population



Randomly Selected N Sample



Hypothesis Testing

- Hypothesis Testing သည် Central Limit Theorem အပေါ် အခြေခံပါသည်။

- ကျွန်တော်တို့သာ Population တစ်ခုရဲ့ Expectation (μ) နှင့် Variance (σ^2)ကိုသာ သိရင် သို့မဟုတ် ခန့်မှန်းနိုင်ရင် Sample များ အတွက် Hypothesis ထုတ်ပြီး မှန်မမှန် စစ်လို့ရပါသည်။

- ပန်းသီးခြံတစ်ခုမှာ ပန်းသီးများ ရှိသည်။ ထိုပန်းသီးများ၏ Weight ရဲ့ Expectation (μ) နှင့် Variance (σ^2)ကိုသိသည် ဆိုပါတော့။ μ = 0.2 kg ဖြစ်ပြီး σ^2= 0.05 ဆိုပါတော့။

- ပန်းသီးခြံရဲ့ အရှေ့ပိုင်းရှိ ပန်းသီးများက အခြားအသီးများ ထက်ပိုလေးသည် ဆိုပြီး Hypothesis ထုတ်သည် ဆိုပါတော့။ ဒါကို Null Hypothesis လို့ခေါ်ပါသည်။ ဒီအဆိုကို မှန်၊ မမှန်စစ်မယ် ဆိုရင်။

- ကျွန်တော်တို့ ပန်းသီးခြံရဲ့ အရှေ့ပိုင်းမှ Random Sample (n) ကို ရွေးရပါမည်။ ထို့နောက် Mean ရှာရပါမည်။

$$Z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

- Z တန်ဖိုး နည်းလေ Null Hypothesis မှန်ဖို့ များလေဖြစ်သည်။ Z တန်ဖိုးသည် သတ်မှတ်ထားသော Threshold Value ထက်နည်းပါက Null Hypothesis မှန်သည်လို့ ယူဆပြီး ပိုများပါက Null Hypothesis မှားသည်ဟု ယူဆပါသည်။

Summary

- ကဲ ကျွန်တော်တို့ Probability နှင့် Statistics ကို ထိုက်သင့်သလောက် သိသွားပါပြီ။
- ဒါဆိုရင် ပထမဆုံး မေးခွန်းကို ဖြေဖို့ အဆင့်သင့် ဖြစ်ပြီလား။ Will you Gamble or Risk?
- အမှန်တော့ ဘဝကြီးသည်ပင် Risk တစ်ခု ဖြစ်ပါသည်။ လုပ်ငန်းလုပ်ခြင်း၊ စာမေးပွဲဖြေခြင်း၊ အိမ်ထောင်ပြုခြင်း ဒါတွေအားလုံးသည် Risk ဖြစ်သည်။ ကျွန်တော်တို့ ဒါတွေကို ကံသေကံမ, မပြောနိုင်ပါ။
- သို့သော် Risk ဖြစ်သည့် အားလျော်စွာ ကျွန်တော်တို့ရဲ့ Probability ကို Increase သို့မဟုတ် Decrease လုပ်လို့ရပါသည်။ ဘယ်အချိန် သေမည်ကို မသိပါ၊ သို့သော် ကျန်းမာရေး ဂရုစိုက်ခြင်းဖြင့် သေဖို့နည်းအောင် လုပ်လို့ရပါသည်။
- Life is not a Gamble; it is a Risk. Thank goodness!