

Lecture 4: Random Variable, Part II

Yi, Yung (이용)

EE210: Probability and Introductory Random Processes KAIST EE

MONTH DAY, 2021

Outline



- Continuous Random Variable
- PDF (Probability Density Function)
- CDF (Cumulative Distribution Function)
- Exponential and Normal Distribution
- Joint PDF, Conditional PDF
- · Bayes' rule for continous and even mixed cases

Roadmap



- Famous discrete random variables used in the community
 - Bernoulli, Uniform, Binomial, Geometric, Poisson, etc.
- Summarizing a random variable: Expectation and Variance
- o Functions of a single random variable, Functions of multiple random variables
- Conditioning for random variables, Independence for random variables
- Continuous random variables
 - Normal, Uniform, Exponential, etc.
- Bayes' rule for random variables





- Many cases when random variable have "continuous values", e.g., velocity of a car



- Many cases when random variable have "continuous values", e.g., velocity of a car

Continuous Random Variable

A rv X is continuous if \exists a function f_X , called probability density function (PDF), s.t.

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



- Many cases when random variable have "continuous values", e.g., velocity of a car

Continuous Random Variable

A rv X is continuous if \exists a function f_X , called probability density function (PDF), s.t.

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

- All of the concepts and methods (expectation, PMFs, and conditioning) for discrete rvs have continuous counterparts



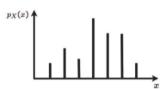
- Many cases when random variable have "continuous values", e.g., velocity of a car

Continuous Random Variable

A rv X is continuous if \exists a function f_X , called probability density function (PDF), s.t.

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

- All of the concepts and methods (expectation, PMFs, and conditioning) for discrete rvs have continuous counterparts



- $\mathbb{P}(a \le X \le b) = \sum_{x:a \le x \le b} p_X(x)$ $p_X(x) \ge 0$, $\sum_x p_X(x) = 1$



- Many cases when random variable have "continuous values", e.g., velocity of a car

Continuous Random Variable

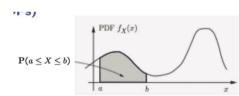
A rv X is continuous if \exists a function f_X , called probability density function (PDF), s.t.

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

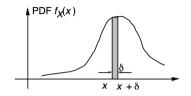
- All of the concepts and methods (expectation, PMFs, and conditioning) for discrete rvs have continuous counterparts



- $\mathbb{P}(a \le X \le b) = \sum_{x:a \le x \le b} p_X(x)$
- $p_X(x) > 0, \sum_{x \in P_X(x)} p_X(x) = 1$

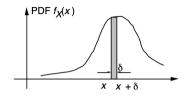


- $\mathbb{P}(a \le X \le b) = \int_a^b f_X(x) dx$ $f_X(x) \ge 0$, $\int_{-\infty}^{\infty} f_X(x) dx = 1$



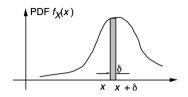
•
$$\mathbb{P}(a \leq X \leq a + \delta) \approx$$





•
$$\mathbb{P}(a \leq X \leq a + \delta) \approx \left| f_X(a) \cdot \delta \right|$$

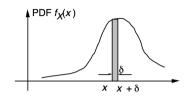




•
$$\mathbb{P}(a \leq X \leq a + \delta) \approx \left| f_X(a) \cdot \delta \right|$$

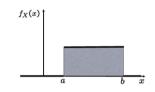
•
$$\mathbb{P}(X = a) = 0$$

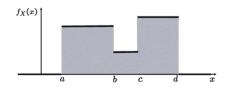




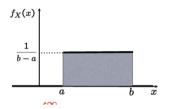
•
$$\mathbb{P}(a \leq X \leq a + \delta) \approx \boxed{f_X(a) \cdot \delta}$$

•
$$\mathbb{P}(X = a) = 0$$



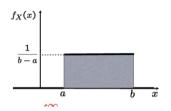






- $\mathbb{E}[X] = \int_{-\infty}^{\infty} x f_X(x) dx =$
- $\mathbb{E}[X^2] = \int_{-\infty}^{\infty} x^2 f_X(x) dx =$
- var[X] =

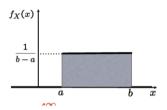




•
$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f_X(x) dx = \int_a^b \frac{x}{b-a} dx = \frac{1}{b-a} \frac{b^2 - a^2}{2} = \frac{b+a}{2}$$

•
$$\mathbb{E}[X^2] = \int_{-\infty}^{\infty} x^2 f_X(x) dx =$$

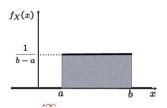




•
$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f_X(x) dx = \int_a^b \frac{x}{b-a} dx = \frac{1}{b-a} \frac{b^2 - a^2}{2} = \frac{b+a}{2}$$

•
$$\mathbb{E}[X^2] = \int_{-\infty}^{\infty} x^2 f_X(x) dx = \int_a^b \frac{x^2}{b-a} dx = \frac{1}{b-a} \frac{b^3 - a^3}{3} = \frac{a^2 + ab + b^2}{3}$$





•
$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f_X(x) dx = \int_a^b \frac{x}{b-a} dx = \frac{1}{b-a} \frac{b^2 - a^2}{2} = \frac{b+a}{2}$$

•
$$\mathbb{E}[X^2] = \int_{-\infty}^{\infty} x^2 f_X(x) dx = \int_a^b \frac{x^2}{b-a} dx = \frac{1}{b-a} \frac{b^3 - a^3}{3} = \frac{a^2 + ab + b^2}{3}$$

•
$$\operatorname{var}[X] = \frac{a^2 + ab + b^2}{3} - \frac{a^2 + 2ab + b^2}{4}$$



• Discrete: PMF, Continuous: PDF



- Discrete: PMF, Continuous: PDF
- Can we describe all rvs with a single mathematical concept?



- Discrete: PMF, Continuous: PDF
- Can we describe all rvs with a single mathematical concept?

$$F_X(x) = \mathbb{P}(X \le x) =$$

$$\begin{cases} \sum_{k \le x} p_X(k), & \text{discrete} \\ \int_{-\infty}^x f_X(t) dt, & \text{continuous} \end{cases}$$



- Discrete: PMF. Continuous: PDF
- Can we describe all rvs with a single mathematical concept?

$$F_X(x) = \mathbb{P}(X \le x) =$$

$$\begin{cases} \sum_{k \le x} p_X(k), & \text{discrete} \\ \int_{-\infty}^x f_X(t) dt, & \text{continuous} \end{cases}$$

• always well defined, because we can always compute the probability for the event $\{X \leq x\}$



- Discrete: PMF. Continuous: PDF
- Can we describe all rvs with a single mathematical concept?

$$F_X(x) = \mathbb{P}(X \le x) =$$

$$\begin{cases} \sum_{k \le x} p_X(k), & \text{discrete} \\ \int_{-\infty}^x f_X(t) dt, & \text{continuous} \end{cases}$$

- always well defined, because we can always compute the probability for the event $\{X \leq x\}$
- CCDF (Complementary CDF): $\mathbb{P}(X > x)$

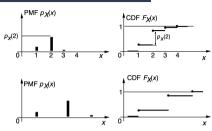


- Discrete: PMF, Continuous: PDF
- Can we describe all rvs with a single mathematical concept?

$$F_X(x) = \mathbb{P}(X \le x) =$$

$$\begin{cases} \sum_{k \le x} p_X(k), & \text{discrete} \\ \int_{-\infty}^x f_X(t) dt, & \text{continuous} \end{cases}$$

- always well defined, because we can always compute the probability for the event $\{X \le x\}$
- CCDF (Complementary CDF): $\mathbb{P}(X > x)$



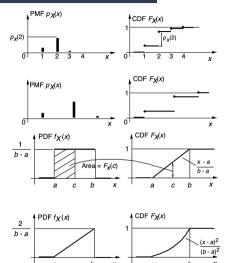


- Discrete: PMF, Continuous: PDF
- Can we describe all rvs with a single mathematical concept?

$$F_X(x) = \mathbb{P}(X \le x) =$$

$$\begin{cases} \sum_{k \le x} p_X(k), & \text{discrete} \\ \int_{-\infty}^x f_X(t) dt, & \text{continuous} \end{cases}$$

- always well defined, because we can always compute the probability for the event $\{X \le x\}$
- CCDF (Complementary CDF): $\mathbb{P}(X > x)$







Non-decreasing



- Non-decreasing
- $F_X(x)$ tends to 1, as $x \to \infty$
- $F_X(x)$ tends to 0, as $x \to -\infty$



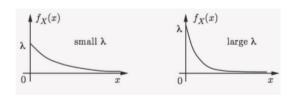
- Non-decreasing
- $F_X(x)$ tends to 1, as $x \to \infty$
- $F_X(x)$ tends to 0, as $x \to -\infty$

Now, let's look at famous continuous random variables popularly used in our life.





$$f_X(x) = egin{cases} \lambda e^{-\lambda x}, & x \geq 0 \ 0, & x < 0 \end{cases} ext{ or } F_X(x) = 1 - e^{-\lambda x}$$

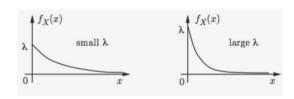




• A rv X is called exponential with λ , if

$$f_X(x) = egin{cases} \lambda e^{-\lambda x}, & x \geq 0 \ 0, & x < 0 \end{cases} ext{ or } F_X(x) = 1 - e^{-\lambda x}$$

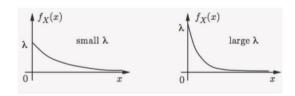
Models a waiting time





$$f_X(x) = egin{cases} \lambda e^{-\lambda x}, & x \geq 0 \ 0, & x < 0 \end{cases} ext{ or } F_X(x) = 1 - e^{-\lambda x}$$

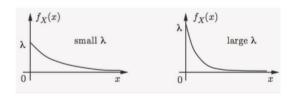
- Models a waiting time
- CCDF $\mathbb{P}(X \geq x) = e^{-\lambda x}$ (waiting time decays exponentially)





$$f_X(x) = egin{cases} \lambda e^{-\lambda x}, & x \geq 0 \ 0, & x < 0 \end{cases} ext{ or } F_X(x) = 1 - e^{-\lambda x}$$

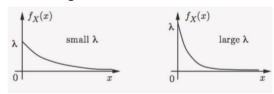
- Models a waiting time
- CCDF $\mathbb{P}(X \ge x) = e^{-\lambda x}$ (waiting time decays exponentially)
- $\mathbb{E}[X] = 1/\lambda$, $\mathbb{E}[X^2] = 2/\lambda^2$, $\text{var}[X] = 1/\lambda^2$





$$f_X(x) = egin{cases} \lambda e^{-\lambda x}, & x \geq 0 \ 0, & x < 0 \end{cases} ext{ or } F_X(x) = 1 - e^{-\lambda x}$$

- Models a waiting time
- CCDF $\mathbb{P}(X \ge x) = e^{-\lambda x}$ (waiting time decays exponentially)
- $\mathbb{E}[X] = 1/\lambda$, $\mathbb{E}[X^2] = 2/\lambda^2$, $\text{var}[X] = 1/\lambda^2$
- (Q) What is the discrete rv which models a waiting time?



Modeling Waiting Time? A Discrete Twin (1)



Modeling Waiting Time? A Discrete Twin (1)



• A discrete twin for modeling waiting times is geometric rvs.



- A discrete twin for modeling waiting times is geometric rvs.
- Models a system evolution over time: Continuous time vs. Discrete time. In many cases, continuous case is the some type of of its corresponding discrete case.



- A discrete twin for modeling waiting times is geometric rvs.
- Models a system evolution over time: Continuous time vs. Discrete time. In many cases, continuous case is the some type of limit of its corresponding discrete case.
- Can you make mathematical description, where geometric and exponential rvs meet each other in the limit?



- A discrete twin for modeling waiting times is geometric rvs.
- Models a system evolution over time: Continuous time vs. Discrete time. In many cases, continuous case is the some type of limit of its corresponding discrete case.
- Can you make mathematical description, where geometric and exponential rvs meet each other in the limit?
- Key idea.
 - Continuous system: Discrete system with



- A discrete twin for modeling waiting times is geometric rvs.
- Models a system evolution over time: Continuous time vs. Discrete time. In many cases, continuous case is the some type of limit of its corresponding discrete case.
- Can you make mathematical description, where geometric and exponential rvs meet each other in the limit?
- Key idea.
 - Continuous system: Discrete system with infinitely many slots whose duration is infinitely small.



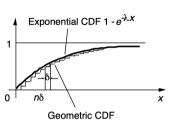
- A discrete twin for modeling waiting times is geometric rvs.
- Models a system evolution over time: Continuous time vs. Discrete time. In many cases, continuous case is the some type of limit of its corresponding discrete case.
- Can you make mathematical description, where geometric and exponential rvs meet each other in the limit?
- Key idea.
 - Continuous system: Discrete system with infinitely many slots whose duration is infinitely small.
- limiting system: $X_{exp}(\lambda)$ with CDF $F_{exp}(\cdot)$



- A discrete twin for modeling waiting times is geometric rvs.
- Models a system evolution over time: Continuous time vs. Discrete time. In many cases, continuous case is the some type of limit of its corresponding discrete case.
- Can you make mathematical description, where geometric and exponential rvs meet each other in the limit?
- Key idea.
 - Continuous system: Discrete system with infinitely many slots whose duration is infinitely small.
- limiting system: $X_{exp}(\lambda)$ with CDF $F_{exp}(\cdot)$
- n-th system: $X^n_{geo}(p_n)$ with CDF $F^n_{geo}(\cdot)$



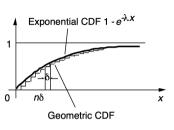
For a given x > 0,





For a given x > 0,

• Define $\delta = \frac{x}{n}$ (a slot length in the *n*-th system)



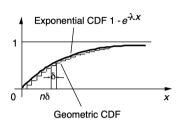


For a given x > 0,

- Define $\delta = \frac{x}{n}$ (a slot length in the *n*-th system)
- Remember

$$F_{exp}(x) = 1 - e^{-\lambda x}$$

 $F_{geo}^{n}(n) = 1 - (1 - p_n)^n$





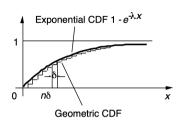
For a given x > 0,

- Define $\delta = \frac{x}{n}$ (a slot length in the *n*-th system)
- Remember

$$F_{exp}(x) = 1 - e^{-\lambda x}$$

 $F_{geo}^{n}(n) = 1 - (1 - p_n)^{n}$

• Choose $p_n = 1 - e^{-\lambda \delta} = 1 - e^{-\lambda \frac{x}{n}}$.





For a given x > 0,

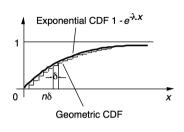
- Define $\delta = \frac{x}{n}$ (a slot length in the *n*-th system)
- Remember

$$F_{exp}(x) = 1 - e^{-\lambda x}$$

$$F_{geo}^{n}(n) = 1 - (1 - p_n)^{n}$$

- Choose $p_n = 1 e^{-\lambda \delta} = 1 e^{-\lambda \frac{x}{n}}$.
- As $n \to \infty$, the slot length $\delta \to 0$ thus $p_n \to 0$
- The CDF values of exponential and *n*-th geometric rvs become equal whenever $x = \delta, 2\delta, 3\delta, \ldots$, i.e.,

$$F_{\text{exp}}(n\delta) = F_{\text{geo}}^n(n), \quad n = 1, 2, \dots$$





For a given x > 0,

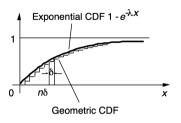
- Define $\delta = \frac{x}{n}$ (a slot length in the *n*-th system)
- Remember

$$F_{exp}(x) = 1 - e^{-\lambda x}$$

 $F_{geo}^{n}(n) = 1 - (1 - p_n)^{n}$

- Choose $p_n = 1 e^{-\lambda \delta} = 1 e^{-\lambda \frac{x}{n}}$.
- As $n \to \infty$, the slot length $\delta \to 0$ thus $p_n \to 0$
- The CDF values of exponential and *n*-th geometric rvs become equal whenever $x = \delta, 2\delta, 3\delta, \ldots$, i.e.,

$$F_{\text{exp}}(n\delta) = F_{\text{geo}}^n(n), \quad n = 1, 2, \dots$$



 As n grows, the number of slots grows, but the success probability over one slot decreases, so that everything is balanced up.



For a given x > 0,

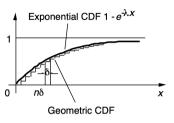
- Define $\delta = \frac{x}{n}$ (a slot length in the *n*-th system)
- Remember

$$F_{exp}(x) = 1 - e^{-\lambda x}$$

 $F_{geo}^{n}(n) = 1 - (1 - p_n)^{n}$

- Choose $p_n = 1 e^{-\lambda \delta} = 1 e^{-\lambda \frac{x}{n}}$.
- As $n \to \infty$, the slot length $\delta \to 0$ thus $p_n \to 0$
- The CDF values of exponential and *n*-th geometric rvs become equal whenever $x = \delta, 2\delta, 3\delta, \ldots$, i.e.,

$$F_{exp}(n\delta) = F_{geo}^n(n), \quad n = 1, 2, \dots$$



- As n grows, the number of slots grows, but the success probability over one slot decreases, so that everything is balanced up.
- As *n* grows, $F_{geo}^n(n)$ approaches $F_{exp}(n\delta)$.

Normal (also called Gaussian) Random Variable



Why important?

- Central limit theorem (중심극한정리)
 - One of the most remarkable findings in the probability theory

Convenient analytical properties

· Modeling aggregate noise with many small, independent noise terms

Normal: PDF, Expectation, Variance



• Standard Normal N(0,1)

$$f_X(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$$

- $\mathbb{E}[X] = 0$
- var[X] = 1

Normal: PDF, Expectation, Variance



• Standard Normal N(0,1)

$$f_X(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$$

• General Normal
$$N(\mu, \sigma^2)$$

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-(x-\mu)^2/2\sigma^2}$$
• $\mathbb{E}[X] = \mu$
• $\operatorname{var}[X] = \sigma^2$

Normal: PDF, Expectation, Variance



• Standard Normal N(0,1)

$$f_X(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$$

- $\mathbb{E}[X] = 0$
- var[X] = 1

Need to check:

- a legitimate PDF or not
- expectation/variance

• General Normal
$$N(\mu, \sigma^2)$$

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-(x-\mu)^2/2\sigma^2}$$
• $\mathbb{E}[X] = \mu$
• $\operatorname{var}[X] = \sigma^2$





• Linear transformation preserves normality

Linear transformation of Normal

If $X \sim \textit{Norm}(\mu, \sigma^2)$, then for $a \neq 0$ and b $Y = aX + b \sim \textit{Norm}(a\mu + b, a^2\sigma^2)$.



• Linear transformation preserves normality

Linear transformation of Normal

If $X \sim \textit{Norm}(\mu, \sigma^2)$, then for $a \neq 0$ and $b \mid Y = aX + b \sim \textit{Norm}(a\mu + b, a^2\sigma^2)$.



• Linear transformation preserves normality

Linear transformation of Normal

If
$$X \sim Norm(\mu, \sigma^2)$$
, then for $a \neq 0$ and $b \mid Y = aX + b \sim Norm(a\mu + b, a^2\sigma^2)$.

• Thus, every normal rv can be standardized :

If
$$X \sim \textit{Norm}(\mu, \sigma^2)$$
, then $Y = \frac{X - \mu}{\sigma} \sim \textit{Norm}(0, 1)$



Linear transformation preserves normality

Linear transformation of Normal

If
$$X \sim Norm(\mu, \sigma^2)$$
, then for $a \neq 0$ and $b \mid Y = aX + b \sim Norm(a\mu + b, a^2\sigma^2)$.

- Thus, every normal rv can be standardized: If $X \sim \textit{Norm}(\mu, \sigma^2)$, then $Y = \frac{\mathsf{X} \mu}{\sigma} \sim \textit{Norm}(0, 1)$
- Thus, we can make the table which records the following CDF values:

$$\Phi(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(Y < y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-t^2/2} dt$$

Example



 Annual snowfall X is modeled as Norm(60, 20²). What is the probability that this year's snowfall is at least 80 inches?

	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09	
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359	
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753	
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141	
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517	
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879	
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224	
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549	
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852	
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133	
0.9	.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	

Example



- Annual snowfall X is modeled as Norm(60, 20²). What is the probability that this year's snowfall is at least 80 inches?
- $Y = \frac{X-60}{20}$.

0.1 S.988 S.638 S.748 S.917 S.957 S.956 S.675 S.714 S.715 0.2 5.793 S.927 S.928 S.931 S.948 S.936 6008 6408 6403 6443 6480 8.918 0.4 6.554 6.950 6.288 6.070 6.762 6762 6648 4.844 8.879 0.5 6.955 6.958 7.019 7.054 7.782 7.747 7.791 7.724 7.747 7.794 7.744 7.746 7.744 7.744 7.747 7.794 7.746 7.744 7.744 7.744 7.744 7.744 7.744 7.744 7.744 7.744 7.744 7.744 7.744 7.744 7.7		.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.2 1.75 3.852 3.871 .910 9.94 .987 .6024 .6044 .610 .6143 .817 .810 .924 .611 .812 .814 .817 .844 .817 .844 .817 .844 .851 .844 .851 .844 .851 .844 .851 .844 .851 .844 .851 .844 .851 .844 .851 .844 .851 .844 .851 .842 .842 .842 .842 .842 .842 .842 .842 .842 .842 .842 .842 .842 .842 .842 .843	0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.3 1.779 9.217 9.225 9.228 9.231 9.828 6.948 4.849 4.879 0.4 4.554 6.859 6.628 6.810 7.702 7.722 7.724 4.844 8.879 0.5 527 7.981 6.828 7.010 7.084 7.123 7.157 7.190 7.224 0.6 7.827 7.911 7.624 7.327 7.701 7.734 7.744 7.734 7.722 7.744 7.734 7.734 7.744 7.744 7.734 7.734 7.744 7.734 7.734 7.744 7.734 7.742 7.744 7.734 7.744 7.734 7.742 7.744 7.744 7.744 7.744 7.744 7.744 7.744 7.744 7.744 7.744 7.744 7.744 7.744 7.744 7.744 7.744 7.744 8.744 8.744 8.744 8.744 8.744 8.744 8.744 8.744 8.744 8.744 8.744	0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.4 0.554 0.591 0.692 0.694 0.700 0.703 0.772 0.808 0.814 0.879 0.5 5.955 0.985 0.084 7.007 .705 .7022 7.121 7.120 7.120 7.120 7.120 7.120 7.120 7.120 7.120 7.120 7.120 7.120 7.120 7.122 7.121 7.121 7.120 7.122 7.121 7.121 7.120 7.222 7.120 7.120 7.124 7.120 7.120 7.124 7.120 <td>0.2</td> <td>.5793</td> <td>.5832</td> <td>.5871</td> <td>.5910</td> <td>.5948</td> <td>.5987</td> <td>.6026</td> <td>.6064</td> <td>.6103</td> <td>.6141</td>	0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
	0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.6 2.527 7.99 7.824 7.93 7.829 7.424 7.43 7.446 7.14 7.446 7.17 7.823<	0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.7 Casto 7511 762 767 7704 7704 782 7826 482 0.9 588 781 789 780 799 782 581 581 8816 813 818 818 818 828 826 828 831 830 836 889 862 888 807 828 826 871 872 874 877 879 869 888 807 822 874 870 870 892 861 880 880 803 892 872 874 870 870 892 962 962 962 962 962 962 962 962 962 962 962 962 97	0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.8 1.881 .7810 7989 .7967 7995 .8023 .80	0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
	0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
1.0	0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
1.1 64.3 86.6 87.6 87.9 87.0 87.0 8.0 83.0 89.0 1.2 84.9 88.8 88.9 89.8 89.8 89.8 99.9 99.2 99.2 99.2 99.2 99.7 90.7 90.7 90.7 90.7 90.7 90.7 90.7 90.7 90.7 90.8<	0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.2 8.49 8.69 9.88 8.907 8.924 8.962 8.88 8.907 8.924 8.962 9.972 9.922 9.926 9.925 9.921 9.922 9.926 9.925 9.922 9.922 9.926 9.925 9.922 9.922 9.926 9.925 9.922 9.926 9.923 9.923 9.923 9.923 9.923 9.923 9.923 9.923 9.923 9.923 9.923 9.923 9.923 9.923 9.924	1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.3 0.802 .904 .908 .908 .908 .908 .908 .908 .911 .9131 .9147 .9162 .9162 .9163 .9163 .9163 .9163 .9163 .9163 .9163 .9163 .9163 .9183 .9184 .9183 .9184 .9183 .9141 .9484 <td>1.1</td> <td>.8643</td> <td>.8665</td> <td>.8686</td> <td>.8708</td> <td>.8729</td> <td>.8749</td> <td>.8770</td> <td>.8790</td> <td>.8810</td> <td>.8830</td>	1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.4 9192 927 922 928 921 926 927 928 936 937 928 936 931 932 932 932 932 932 932 932 932 932 932 932 932 932 932 942 942 942 943 942 942 942 942 942 942 942 942 942 942 942 942 942 942 942 942 942 943 942 943 943 942 943 943 943 943 943 943 943 943 943 943 943 943 943 943 943 943 943 943 943 944 943 944 943 944 <td>1.2</td> <td>.8849</td> <td>.8869</td> <td>.8888</td> <td>.8907</td> <td>.8925</td> <td>.8944</td> <td>.8962</td> <td>.8980</td> <td>.8997</td> <td>.9015</td>	1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.0		.9032	.9049	.9066	.9082	.9099	.9115		.9147	.9162	.9177
1.6 9452 9463 9474 9484 9495 9005 9151 9325 9353 9362 1.8 941 9619 9656 964 9671 9678 9686 9682 9682 9683 9699 9761 9762 9811 9843 9843 9843 9843 9843 9843 9843 9843 9843 9843 9843 9843 9843 9843 9843 9843 </td <td>1.4</td> <td>.9192</td> <td>.9207</td> <td>.9222</td> <td>.9236</td> <td>.9251</td> <td>.9265</td> <td>.9279</td> <td>.9292</td> <td>.9306</td> <td>.9319</td>	1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.7 5.54 9.64 9.75 9.82 9.09 9.09 9.016 6.62 9.62 9.75 1.8 9.41 9.75 9.73 9.73 9.74 9.76 9.77 9.7	1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.8 9641 9649 9656 9644 9671 9678 9680 9680 9680 9680 9680 9680 9680 9681 9676 9767 9767 9767 9767 9767 9761 9767 9761 9767 9761 9767 9761 9767 9761 9767 9761 9767 9761 9767 9761 9762 9812		.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.9 9713 9719 9766 9782 9783 9744 9750 9766 9761 9671 2.0 9272 9778 9783 9788 9793 9808 9818 9861 9868 9814 9846 9868 9814 9846 9863 9841 9864 9863 9812 9844 9873 9801 9913 9913 9913 9913 9913 9913 9913 9913 9913 9913 9913 9914 9943 9944 9842 9842 9842 9844 9872 9924	1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
20 19772 9778 9788 9788 9780 9802 9801 9812 8812	1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
2.1 18/21 19/22 98/20 98/30 98/34 98/35 98/42 98/35 98/34 98/35 98/34 98/35 98/34 98/35 98/34 98/35 98/34 98/35 9	1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.2 9861 9864 9888 9871 9878 9881 9881 9891 991 991 9910 9910 9910 9913 9913 9913 9913 9913 9913 9913 9913 9913 9933 9934 99	2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.3 J883 J996 J898 J901 J906 J900 J901 J913 J913 J914 J914 J918 J918 J913 J914 J914 J918 J918 J913 J914	2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.4 9918 9920 9922 9925 9927 9929 9831 9832 9934 9934 2.5 5938 940 .9941 .9945 .9945 .9949 .9961 .9962 .9961 .9962 .9961 .9962 .9961 .9962 .9963 .9964 .9948 .9949 .9961 .9962 .9963 .9964 .9949 .9961 .9962 .9963 .9964 .9949 .9961 .9962 .9963 .9964 .9949 .9962 .9963 .9964 .9949 .9972 .9972 .9973 .9974 .9974 .9974 .9972 .9973 .9974 <	2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.5 .9938 .9940 .9941 .9943 .9945 .9946 .9948 .9949 .9951 .9952 2.6 .9963 .9956 .9957 .9969 .9960 .9961 .9962 .9963 .9968 .9969 .9970 .9971 .9978 .9964 .9969 .9970 .9971 .9978 .9974 .9978 .9974 .9978 .9979 .9978 .9978 .9980 .9981 .9981 .9981 .9981 .9981 .9981 .9981 .9981 .9981 .9981 .9981 .9981 .9981 .9982 .9983 .9983 .9983 .9983 .9984 .9979 .9979 .9979 .9978 .9984 .9979 .9984 </td <td>2.3</td> <td>.9893</td> <td>.9896</td> <td>.9898</td> <td>.9901</td> <td>.9904</td> <td>.9906</td> <td>.9909</td> <td>.9911</td> <td>.9913</td> <td>.9916</td>	2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
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	2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.7 .9965 .9966 .9967 .9988 .9969 .9970 .9971 .9972 .9973 .9974 2.8 .9974 .9975 .9976 .9977 .9977 .9978 .9979 .9979 .9980 .9981	2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.8 .9974 .9975 .9976 .9977 .9978 .9979 .9979 .9980 .9981	2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
	2.7	.9965	.9966	.9967	.9968	.9969		.9971	.9972	.9973	.9974
2.9 .9981 .9982 .9982 .9983 .9984 .9984 .9985 .9985 .9986 .9986	2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
	2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986

Example^l



- Annual snowfall X is modeled as Norm(60, 20²). What is the probability that this year's snowfall is at least 80 inches?
- $Y = \frac{X-60}{20}$. $\mathbb{P}(X \ge 80) = \mathbb{P}(Y \ge \frac{80-60}{20})$ $= \mathbb{P}(Y \ge 1) = 1 - \Phi(1)$ = 1 - 0.8413 = 0.1587

	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.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	.9543
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9708
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	.996
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

Roadmap



- Famous discrete random variables used in the community
 - Bernoulli, Uniform, Binomial, Geometric, Poisson, etc.
- Summarizing a random variable: Expectation and Variance
- Functions of a single random variable, Functions of multiple random variables
- Conditioning for random variables, Independence for random variables
- Continuous random variables
 - Normal, Uniform, Exponential, etc.
- Bayes' rule for random variables
- ** Continuous counterparts are intuitively understandable. So, we will be quick at reviewing them.



Jointly Continuous

Two continuous rvs are if a non-negative function $f_{X,Y}(x,y)$

(called joint PDF) satisfies: for $\boxed{\text{every}}$ subset B of the two dimensional plane,

$$\mathbb{P}((X,Y)\in B)=\iint_{(x,y)\in B}f_{X,Y}(x,y)dxdy$$



Jointly Continuous

Two continuous rvs are jointly continuous if a non-negative function $f_{X,Y}(x,y)$ (called joint PDF) satisfies: for every subset B of the two dimensional plane,

$$\mathbb{P}((X,Y)\in B)=\iint_{(x,y)\in B}f_{X,Y}(x,y)dxdy$$



Jointly Continuous

Two continuous rvs are jointly continuous if a non-negative function $f_{X,Y}(x,y)$ (called joint PDF) satisfies: for every subset B of the two dimensional plane,

$$\mathbb{P}((X,Y)\in B)=\iint_{(x,y)\in B}f_{X,Y}(x,y)dxdy$$

1. The joint PDF is used to calculate probabilities

$$\mathbb{P}((X,Y)\in B)=\iint_{(x,y)\in B}f_{X,Y}(x,y)dxdy$$

Our particular interest: $B = \{(x, y) \mid a \le x \le b, c \le y \le d\}$





2. The marginal PDFs of X and Y are from the joint PDF as:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dy, \quad f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dx$$



2. The marginal PDFs of X and Y are from the joint PDF as:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dy, \quad f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dx$$

3. The joint CDF is defined by $F_{X,Y}(x,y) = \mathbb{P}(X \le x, Y \le y)$, and determines the joint PDF as:

$$f_{X,Y}(x,y) = \frac{\partial^2 F_{x,y}}{\partial x \partial y}(x,y)$$



2. The marginal PDFs of X and Y are from the joint PDF as:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dy, \quad f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dx$$

3. The joint CDF is defined by $F_{X,Y}(x,y) = \mathbb{P}(X \le x, Y \le y)$, and determines the joint PDF as:

$$f_{X,Y}(x,y) = \frac{\partial^2 F_{x,y}}{\partial x \partial y}(x,y)$$

4. A function g(X, Y) of X and Y defines a new random variable, and

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

Continuous: Conditional PDF given an event



* Conditional PDF, given an event

* Conditional PDF, given $X \in B$

Continuous: Conditional PDF given an event



- * Conditional PDF, given an event
- $f_X(x) \cdot \delta \approx \mathbb{P}(x \le X \le x + \delta)$ $f_{X|A}(x) \cdot \delta \approx \mathbb{P}(x \le X \le x + \delta|A)$

* Conditional PDF, given $X \in B$

Continuous: Conditional PDF given an event



- * Conditional PDF, given an event
- $f_X(x) \cdot \delta \approx \mathbb{P}(x \le X \le x + \delta)$ $f_{X|A}(x) \cdot \delta \approx \mathbb{P}(x \le X \le x + \delta|A)$
- $\mathbb{P}(X \in B) = \int_B f_X(x) dx$ $\mathbb{P}(X \in B|A) = \int_B f_{X|A}(x) dx$

Note: A is an event, but B is a subset that includes the possible values which can be taken by the rv X.

* Conditional PDF, given $X \in B$

Continuous: Conditional PDF given an event



- * Conditional PDF, given an event
- $f_X(x) \cdot \delta \approx \mathbb{P}(x \le X \le x + \delta)$ $f_{X|A}(x) \cdot \delta \approx \mathbb{P}(x \le X \le x + \delta|A)$
- $\mathbb{P}(X \in B) = \int_B f_X(x) dx$ $\mathbb{P}(X \in B|A) = \int_B f_{X|A}(x) dx$

Note: A is an event, but B is a subset that includes the possible values which can be taken by the rv X.

•
$$\int f_{X|A}(x) = 1$$

* Conditional PDF, given $X \in B$

Continuous: Conditional PDF given an event



- * Conditional PDF, given an event
- $f_X(x) \cdot \delta \approx \mathbb{P}(x \le X \le x + \delta)$ $f_{X|A}(x) \cdot \delta \approx \mathbb{P}(x \le X \le x + \delta|A)$
- $\mathbb{P}(X \in B) = \int_B f_X(x) dx$ $\mathbb{P}(X \in B|A) = \int_B f_{X|A}(x) dx$

Note: A is an event, but B is a subset that includes the possible values which can be taken by the rv X.

• $\int f_{X|A}(x) = 1$

* Conditional PDF, given $X \in B$

$$\mathbb{P}(x \le X \le x + \delta | X \in B) \approx f_{X|X \in B}(x) \cdot \delta$$

$$f_{X|X\in B}(x) = \begin{cases} 0, & \text{if } x \notin B \\ \frac{f_X(x)}{\mathbb{P}(B)}, & \text{if } x \in B \end{cases}$$

Continuous: Conditional PDF given an event



- * Conditional PDF, given an event
- $f_X(x) \cdot \delta \approx \mathbb{P}(x \le X \le x + \delta)$ $f_{X|A}(x) \cdot \delta \approx \mathbb{P}(x \le X \le x + \delta|A)$
- $\mathbb{P}(X \in B) = \int_B f_X(x) dx$ $\mathbb{P}(X \in B|A) = \int_B f_{X|A}(x) dx$

Note: A is an event, but B is a subset that includes the possible values which can be taken by the rv X.

• $\int f_{X|A}(x) = 1$

* Conditional PDF, given $X \in B$

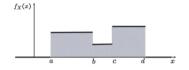
$$\mathbb{P}(x \le X \le x + \delta | X \in B) \approx f_{X|X \in B}(x) \cdot \delta$$

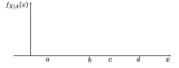
$$f_{X|X\in B}(x) = \begin{cases} 0, & \text{if } x \notin B \\ \frac{f_X(x)}{\mathbb{P}(B)}, & \text{if } x \in B \end{cases}$$

(Q) In the discrete, we consider the event $\{X = x\}$, not $\{X \in B\}$. Why?



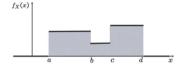
$$A = \left\{ \frac{a+b}{2} \le X \le b \right\}$$

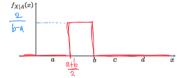






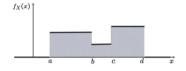
$$A = \left\{ \frac{a+b}{2} \le X \le b \right\}$$

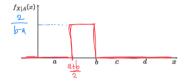






$$A = \left\{ \frac{a+b}{2} \le X \le b \right\}$$



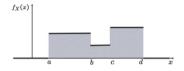


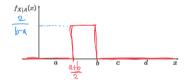
•
$$\mathbb{E}[X] = \int x f_X(x) dx$$

 $\mathbb{E}[X|A] = \int x f_{X|A}(x) dx$



$$A = \left\{ \frac{a+b}{2} \le X \le b \right\}$$





•
$$\mathbb{E}[X] = \int x f_X(x) dx$$

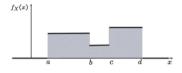
 $\mathbb{E}[X|A] = \int x f_{X|A}(x) dx$

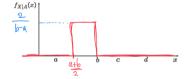
•
$$\mathbb{E}[g(X)] = \int g(x) f_X(x) dx$$

 $\mathbb{E}[g(X)|A] = \int g(x) f_{X|A}(x) dx$



$$A = \left\{ \frac{a+b}{2} \le X \le b \right\}$$





•
$$\mathbb{E}[X] = \int x f_X(x) dx$$

 $\mathbb{E}[X|A] = \int x f_{X|A}(x) dx$

•
$$\mathbb{E}[g(X)] = \int g(x) f_X(x) dx$$

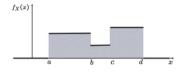
 $\mathbb{E}[g(X)|A] = \int g(x) f_{X|A}(x) dx$

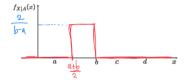
$$\mathbb{E}[X|A] =$$

$$\mathbb{E}[X^2|A] =$$



$$A = \left\{ \frac{a+b}{2} \le X \le b \right\}$$





•
$$\mathbb{E}[X] = \int x f_X(x) dx$$

 $\mathbb{E}[X|A] = \int x f_{X|A}(x) dx$

•
$$\mathbb{E}[g(X)] = \int g(x) f_X(x) dx$$

 $\mathbb{E}[g(X)|A] = \int g(x) f_{X|A}(x) dx$

$$\mathbb{E}[X|A] = \int_{(a+b)/2}^{b} x \frac{2}{b-a} dx = \frac{a}{4} + \frac{3b}{4}$$

$$\mathbb{E}[X^{2}|A] = \int_{(a+b)/2}^{b} x^{2} \frac{2}{b-a} dx =$$



• Exponential rv is a continous counterpart of geometric rv.



- Exponential rv is a continous counterpart of geometric rv.
- Thus, expected to be memeoryless.



- Exponential rv is a continous counterpart of geometric rv.
- Thus, expected to be memeoryless.

Definition. A random variable X is called memoryless if, for any $n, m \ge 0$,

$$\mathbb{P}(X > n + m | X > m) = \mathbb{P}(X > n)$$



- Exponential rv is a continous counterpart of geometric rv.
- Thus, expected to be memeoryless.

Definition. A random variable X is called memoryless if, for any $n, m \ge 0$, $\mathbb{P}(X > n + m | X > m) = \mathbb{P}(X > n)$

• Proof. Note that $\mathbb{P}(X > x) = e^{-\lambda x}$.



- Exponential rv is a continous counterpart of geometric rv.
- Thus, expected to be memeoryless.

Definition. A random variable X is called memoryless if, for any $n, m \ge 0$, $\mathbb{P}(X > n + m | X > m) = \mathbb{P}(X > n)$

• Proof. Note that $\mathbb{P}(X > x) = e^{-\lambda x}$. Then,

$$\mathbb{P}(X>n+m|X>m)=\frac{\mathbb{P}(X>n+m)}{\mathbb{P}(X>m)}=\frac{e^{-\lambda(n+m)}}{e^{-\lambda m}}=e^{-\lambda n}=\mathbb{P}(X>n)$$



Partition of Ω into A_1, A_2, A_3, \dots

* Discrete case

* Continuous case



Partition of Ω into A_1, A_2, A_3, \dots

* Discrete case

Total Probability Theorem

$$p_X(x) = \sum_i \mathbb{P}(A_i)\mathbb{P}(X = x|A_i)$$

= $\sum_i \mathbb{P}(A_i)p_{X|A_i}(x)$

Total Expectation Theorem

$$\mathbb{E}[X] = \sum_{i} \mathbb{P}(A_i) \mathbb{E}[X|A_i]$$

* Continuous case



Partition of Ω into A_1, A_2, A_3, \dots

* Discrete case

Total Probability Theorem

$$\rho_X(x) = \sum_i \mathbb{P}(A_i) \mathbb{P}(X = x | A_i)$$

$$= \sum_i \mathbb{P}(A_i) \rho_{X|A_i}(x)$$

Total Expectation Theorem

$$\mathbb{E}[X] = \sum_{i} \mathbb{P}(A_i) \mathbb{E}[X|A_i]$$

* Continuous case

Total Probability Theorem

$$f_X(x) = \sum_i \mathbb{P}(A_i) f_{X|A_i}(x)$$



Partition of Ω into A_1, A_2, A_3, \dots

* Discrete case

Total Probability Theorem

$$p_X(x) = \sum_i \mathbb{P}(A_i)\mathbb{P}(X = x|A_i)$$

= $\sum_i \mathbb{P}(A_i)p_{X|A_i}(x)$

Total Expectation Theorem

$$\mathbb{E}[X] = \sum_i \mathbb{P}(A_i) \mathbb{E}[X|A_i]$$

* Continuous case

Total Probability Theorem

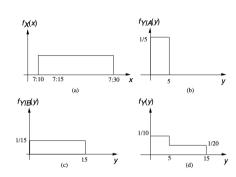
$$f_X(x) = \sum_i \mathbb{P}(A_i) f_{X|A_i}(x)$$

Total Expectation Theorem

$$\mathbb{E}[X] = \sum_{i} \mathbb{P}(A_i) \mathbb{E}[X|A_i]$$

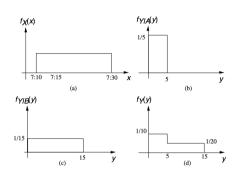


- Your train's arrival every quarter hour (0, 15min, 30min, 45min).
- Your arrival \sim uniform(7:10, 7:30) am.



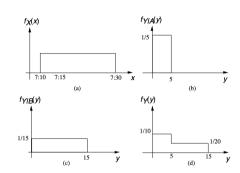


- Your train's arrival every quarter hour (0, 15min, 30min, 45min).
- Your arrival \sim uniform(7:10, 7:30) am.
- What is the PDF of waiting time for the first train?





- Your train's arrival every quarter hour (0, 15min, 30min, 45min).
- Your arrival \sim uniform(7:10, 7:30) am.
- What is the PDF of waiting time for the first train?
- X: your arrival time, Y: waiting time.

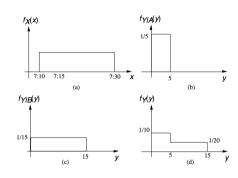




- Your train's arrival every quarter hour (0, 15min, 30min, 45min).
- Your arrival \sim uniform(7:10, 7:30) am.
- What is the PDF of waiting time for the first train?
- X : your arrival time, Y : waiting time.
- The value of X makes a different waiting time. So, consider two events:

$$A = \{7:10 \le X \le 7:15\}$$

$$B = \{7:15 \le X \le 7:30\}$$

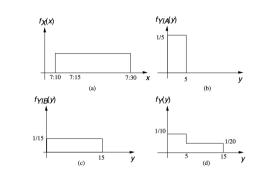




- Your train's arrival every quarter hour (0, 15min, 30min, 45min).
- Your arrival \sim uniform(7:10, 7:30) am.
- What is the PDF of waiting time for the first train?
- X: your arrival time, Y: waiting time.
- The value of X makes a different waiting time. So, consider two events:

$$A = \{7:10 \le X \le 7:15\}$$

$$B = \{7:15 \le X \le 7:30\}$$



$$f_Y(y) = \mathbb{P}(A)f_{Y|A}(y) + \mathbb{P}(B)f_{Y|B}(y)$$
 for $0 \le y \le 5$

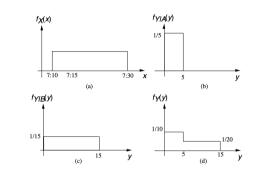
for
$$5 < y \le 15$$



- Your train's arrival every quarter hour (0, 15min, 30min, 45min).
- Your arrival \sim uniform(7:10, 7:30) am.
- What is the PDF of waiting time for the first train?
- X: your arrival time, Y: waiting time.
- The value of X makes a different waiting time. So, consider two events:

$$A = \{7:10 \le X \le 7:15\}$$

$$B = \{7:15 \le X \le 7:30\}$$



$$f_{Y}(y) = \mathbb{P}(A)f_{Y|A}(y) + \mathbb{P}(B)f_{Y|B}(y)$$

$$f_{Y}(y) = \frac{1}{4}\frac{1}{5} + \frac{3}{4}\frac{1}{15} = \frac{1}{10}, \text{ for } 0 \le y \le 5$$

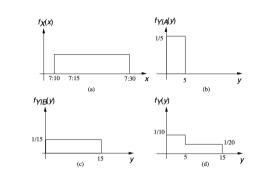
for
$$5 < y \le 15$$



- Your train's arrival every quarter hour (0, 15min, 30min, 45min).
- Your arrival \sim uniform(7:10, 7:30) am.
- What is the PDF of waiting time for the first train?
- X: your arrival time, Y: waiting time.
- The value of X makes a different waiting time. So, consider two events:

$$A = \{7:10 \le X \le 7:15\}$$

$$B = \{7:15 \le X \le 7:30\}$$



$$f_Y(y) = \mathbb{P}(A)f_{Y|A}(y) + \mathbb{P}(B)f_{Y|B}(y)$$

$$f_Y(y) = \frac{1}{4}\frac{1}{5} + \frac{3}{4}\frac{1}{15} = \frac{1}{10}, \quad \text{for } 0 \le y \le 5$$

$$f_Y(y) = \frac{1}{4}0 + \frac{3}{4}\frac{1}{15} = \frac{1}{20}, \quad \text{for } 5 < y \le 15$$



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



- $p_{X|Y}(x|y) = \frac{p_{X,Y}(x,y)}{p_{Y}(y)}$
- Similarly, for $f_Y(y) > 0$,

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



- $p_{X|Y}(x|y) = \frac{p_{X,Y}(x,y)}{p_{Y}(y)}$
- Similarly, for $f_Y(y) > 0$,

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

• Remember: For a fixed event A, $\mathbb{P}(\cdot|A)$ is a legitimate probability law.



- $p_{X|Y}(x|y) = \frac{p_{X,Y}(x,y)}{p_{Y}(y)}$
- Similarly, for $f_Y(y) > 0$,

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

- Remember: For a fixed event A, P(·|A) is a legitimate probability law.
- Similarly, For a fixed y, $f_{X|Y}(x|y)$ is a legitimate PDF, since

$$\int_{-\infty}^{\infty} f_{X|Y}(x|y) \frac{1}{dx} = \frac{\int_{-\infty}^{\infty} f_{X,Y}(x,y) dx}{f_{Y}(y)} = 1$$



- $p_{X|Y}(x|y) = \frac{p_{X,Y}(x,y)}{p_{Y}(y)}$
- Similarly, for $f_Y(y) > 0$,

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

- Remember: For a fixed event A, P(·|A) is a legitimate probability law.
- Similarly, For a fixed y, $f_{X|Y}(x|y)$ is a legitimate PDF, since

$$\int_{-\infty}^{\infty} f_{X|Y}(x|y) \frac{dx}{dx} = \frac{\int_{-\infty}^{\infty} f_{X,Y}(x,y) dx}{f_{Y}(y)} = 1$$

Multiplication rule.

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

= $f_X(x)f_{Y|X}(y|x)$



- $p_{X|Y}(x|y) = \frac{p_{X,Y}(x,y)}{p_{Y}(y)}$
- Similarly, for $f_Y(y) > 0$,

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

- Remember: For a fixed event A, P(·|A) is a legitimate probability law.
- Similarly, For a fixed y, $f_{X|Y}(x|y)$ is a legitimate PDF, since

$$\int_{-\infty}^{\infty} f_{X|Y}(x|y) \frac{dx}{dx} = \frac{\int_{-\infty}^{\infty} f_{X,Y}(x,y) dx}{f_{Y}(y)} = 1$$

Multiplication rule.

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

= $f_X(x)f_{Y|X}(y|x)$

• Total prob./exp. theorem.

$$f_X(x) = \int_{-\infty}^{\infty} f_Y(y) f_{X|Y}(x|y) dy$$

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

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



- $p_{X|Y}(x|y) = \frac{p_{X,Y}(x,y)}{p_{Y}(y)}$
- Similarly, for $f_Y(y) > 0$,

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

- Remember: For a fixed event A, P(·|A) is a legitimate probability law.
- Similarly, For a fixed y, $f_{X|Y}(x|y)$ is a legitimate PDF, since

$$\int_{-\infty}^{\infty} f_{X|Y}(x|y) \frac{dx}{dx} = \frac{\int_{-\infty}^{\infty} f_{X,Y}(x,y) dx}{f_{Y}(y)} = 1$$

Multiplication rule.

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

= $f_X(x)f_{Y|X}(y|x)$

• Total prob./exp. theorem.

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

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

• Independence.

$$f_{X,Y}(x,y) = f_X(x)f_Y(y)$$
, for all x and y



- Break a stick of length / twice
 - first break at $X \sim uniform[0.1]$
 - second break at $Y \sim \textit{uniform}[0, X]$



- Break a stick of length / twice
 - first break at $X \sim uniform[0.1]$
 - second break at $Y \sim \textit{uniform}[0, X]$
- (Q) What is $\mathbb{E}[Y]$?



- Break a stick of length / twice
 - first break at $X \sim uniform[0.1]$
 - second break at $Y \sim uniform[0, X]$
- (Q) What is $\mathbb{E}[Y]$?
- Since Y depends on X, the total expectation theorem seems useful.

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



- Break a stick of length / twice
 - first break at $X \sim uniform[0.1]$
 - second break at $Y \sim uniform[0, X]$
- (Q) What is $\mathbb{E}[Y]$?
- Since Y depends on X, the total expectation theorem seems useful.

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

Using the TET,

$$\mathbb{E}[Y] = \int_0^I \frac{1}{I} \mathbb{E}[Y|X = x] dx$$
$$= \int_0^I \frac{1}{I} \frac{x}{2} dx = \frac{I}{4}$$

Example: Stick-breaking (Ch 3. Prob 21)



- Break a stick of length / twice
 - first break at $X \sim uniform[0.1]$
 - second break at $Y \sim uniform[0, X]$
- (Q) What is $\mathbb{E}[Y]$?
- Since Y depends on X, the total expectation theorem seems useful.

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

Using the TET,

$$\mathbb{E}[Y] = \int_0^I \frac{1}{I} \mathbb{E}[Y|X = x] dx$$
$$= \int_0^I \frac{1}{I} \frac{x}{2} dx = \frac{I}{4}$$

• $f_X(x)$ and $f_{Y|X}(y|x)$ seems easy to compute. Thus,

$$f_{X,Y}(x,y) = f_X(x)f_{Y|X}(y|x) = \frac{1}{I} \cdot \frac{1}{x}$$

You can do many other things with the joint PDF.

Roadmap



- Famous discrete random variables used in the community
 - Bernoulli, Uniform, Binomial, Geometric, Poisson, etc.
- Summarizing a random variable: Expectation and Variance
- o Functions of a single random variable, Functions of multiple random variables
- Conditioning for random variables, Independence for random variables
- Continuous random variables
 - Normal, Uniform, Exponential, etc.
- Bayes' rule for random variables

Bayes Rule for Continuous



- X: state/cause/original value $\rightarrow Y$: result/resulting action/noisy measurement
- Model: $\mathbb{P}(X)$ (prior) and $\mathbb{P}(Y|X)$ (cause \to result)
- Inference: $\mathbb{P}(X|Y)$?

Bayes Rule for Continuous



- X: state/cause/original value $\rightarrow Y$: result/resulting action/noisy measurement
- Model: $\mathbb{P}(X)$ (prior) and $\mathbb{P}(Y|X)$ (cause \to result)
- Inference: $\mathbb{P}(X|Y)$?

$$p_{X,Y}(x,y) = p_X(x)p_{Y|X}(y|x)$$

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

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

$$p_Y(y) = \sum_{x'} p_X(x')p_{Y|X}(y|x')$$

$$f_{X,Y}(x,y) = f_X(x)f_{Y|X}(y|x)$$

$$= f_Y(y)f_{X|Y}(x|y)$$

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

$$f_Y(y) = \int f_X(x')f_{Y|X}(y|x')dx'$$



K: discrete, Y: continuous



K: discrete, Y: continuous

• Inference of K given Y

• Inference of Y given K



K: discrete, Y: continuous

• Inference of *K* given *Y*

$$p_{K|Y}(k|y) = \frac{p_K(k)f_{Y|K}(y|k)}{f_Y(y)}$$
$$f_Y(y) = \sum_{k'} p_K(k')f_{Y|K}(y|k')$$

• Inference of *Y* given *K*



K: discrete, Y: continuous

• Inference of K given Y

$$p_{K|Y}(k|y) = \frac{p_K(k)f_{Y|K}(y|k)}{f_Y(y)}$$
$$f_Y(y) = \sum_{k'} p_K(k')f_{Y|K}(y|k')$$

• Inference of Y given K

$$f_{Y|K}(y|k) = \frac{f_Y(y)p_{K|Y}(k|y)}{p_K(k)}$$
$$p_K(k) = \int f_Y(y')p_{K|Y}(k|y')dy'$$



$$p_{K|Y}(k|y) = \frac{p_K(k)f_{Y|K}(y|k)}{f_Y(y)}, \quad f_Y(y) = \sum_{k'} p_K(k')f_{Y|K}(y|k')$$



Inference of discrete *K* given continuous *Y*:

$$p_{K|Y}(k|y) = \frac{p_K(k)f_{Y|K}(y|k)}{f_Y(y)}, \quad f_Y(y) = \sum_{k'} p_K(k')f_{Y|K}(y|k')$$

• K: -1, +1, original signal, equally likely. $p_K(1) = 1/2$, $p_K(-1) = 1/2$.



$$p_{K|Y}(k|y) = \frac{p_K(k)f_{Y|K}(y|k)}{f_Y(y)}, \quad f_Y(y) = \sum_{k'} p_K(k')f_{Y|K}(y|k')$$

- K: -1, +1, original signal, equally likely. $p_K(1) = 1/2$, $p_K(-1) = 1/2$.
- Y: measured signal with Gaussian noise, $Y = K + W, \ W \sim N(0,1)$



$$p_{K|Y}(k|y) = \frac{p_K(k)f_{Y|K}(y|k)}{f_Y(y)}, \quad f_Y(y) = \sum_{k'} p_K(k')f_{Y|K}(y|k')$$

- K: -1, +1, original signal, equally likely. $p_K(1) = 1/2, p_K(-1) = 1/2$.
- Y: measured signal with Gaussian noise, Y = K + W, $W \sim N(0,1)$
- Your received signal = 0.7. What's your guess about the original signal?



$$p_{K|Y}(k|y) = \frac{p_K(k)f_{Y|K}(y|k)}{f_Y(y)}, \quad f_Y(y) = \sum_{k'} p_K(k')f_{Y|K}(y|k')$$

- K: -1, +1, original signal, equally likely. $p_K(1) = 1/2$, $p_K(-1) = 1/2$.
- Y: measured signal with Gaussian noise, Y = K + W, $W \sim N(0,1)$
- Your received signal = 0.7. What's your guess about the original signal?
- Your received signal = -0.2. What's your guess about the original signal?



$$p_{K|Y}(k|y) = \frac{p_K(k)f_{Y|K}(y|k)}{f_Y(y)}, \quad f_Y(y) = \sum_{k'} p_K(k')f_{Y|K}(y|k')$$

- K: -1, +1, original signal, equally likely. $p_K(1) = 1/2$, $p_K(-1) = 1/2$.
- Y: measured signal with Gaussian noise, Y = K + W, $W \sim N(0,1)$
- Your received signal = 0.7. What's your guess about the original signal? +1
- Your received signal = -0.2. What's your guess about the original signal? -1



• $Y|K = 1 \sim N(1,1)$ and $Y|K = -1 \sim N(-1,1)$.



• $Y|K = 1 \sim N(1,1)$ and $Y|K = -1 \sim N(-1,1)$.

$$f_{Y|K}(y|k) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(y-k)^2}, \quad k = 1, -1$$



• $Y|K = 1 \sim N(1,1)$ and $Y|K = -1 \sim N(-1,1)$.

$$f_{Y|K}(y|k) = rac{1}{\sqrt{2\pi}}e^{-rac{1}{2}(y-k)^2}, \quad k = 1, -1$$
 $f_{Y}(y) = rac{1}{2}rac{1}{\sqrt{2\pi}}e^{-rac{1}{2}(y+1)^2} + rac{1}{2}rac{1}{\sqrt{2\pi}}e^{-rac{1}{2}(y-1)^2}$



• $Y|K = 1 \sim N(1,1)$ and $Y|K = -1 \sim N(-1,1)$.

$$f_{Y|K}(y|k) = rac{1}{\sqrt{2\pi}}e^{-rac{1}{2}(y-k)^2}, \quad k = 1, -1$$
 $f_{Y}(y) = rac{1}{2}rac{1}{\sqrt{2\pi}}e^{-rac{1}{2}(y+1)^2} + rac{1}{2}rac{1}{\sqrt{2\pi}}e^{-rac{1}{2}(y-1)^2}$

• Probability that K = 1, given Y = y? After some algebra,

$$ho_{K|Y}(1|y) = rac{1}{1 + e^{-2y}}$$

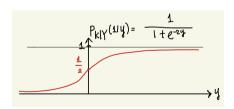


• $Y|K = 1 \sim N(1,1)$ and $Y|K = -1 \sim N(-1,1)$.

$$f_{Y|K}(y|k) = rac{1}{\sqrt{2\pi}}e^{-rac{1}{2}(y-k)^2}, \quad k = 1, -1$$
 $f_{Y}(y) = rac{1}{2}rac{1}{\sqrt{2\pi}}e^{-rac{1}{2}(y+1)^2} + rac{1}{2}rac{1}{\sqrt{2\pi}}e^{-rac{1}{2}(y-1)^2}$

• Probability that K = 1, given Y = y? After some algebra,

$$ho_{K|Y}(1|y) = rac{1}{1 + e^{-2y}}$$





Questions?

Review Questions



- 1) What is PDF and CDF?
- 2) Why do we need CDF?
- 3) What are joint/marginal/conditional PDFs?
- 4) Explain memorylessness of exponential random variables.
- Explain the version of Bayes' rule for continuous and mixed random variables.