MA-120 Probability and Statistics

Nazar Khan

PUCIT

Lecture 11: Probability Distributions - Continuous

Outline

1. Probability Density vs. Probability Mass
2. Normal Density -- Queen of densities
3. Standard Normal Density
4. Standardization
   – z-score = amount of standard deviations away from the mean
5. Standard Normal Table
6. Normal approximation to discrete densities
   – Continuity correction
   – Normal approximation of Binomial density

So far we have covered …

1. Random Experiments – processes with uncertain outcomes
2. Sample Space – outcomes of experiments
3. Events
4. Probability – assigns numbers between 0 and 1 to events
5. Independence – $P(ABC\ldots)=P(A)P(B)P(C)\ldots$

So far we have covered …

6. Random Variables – assign labels to each outcome
   - X(HHH)=3 if random variable X is the number of heads
   - X(HHH)=0 if random variable X is the number of tails
7. Probability **Density** of a random variable

Values of X	0	1	2	3	4
Probabilities	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{1}{16}$

← **Labels**

← **Probabilities**

8. Cumulative Probability **Distribution** of a random variable – P(*X*<=*t*)

So far we have covered …

- **Discrete** random variables – set of outputs is real valued, <u>countable</u> set
- Now we study **continuous** random variables
  - set of outputs is real valued, <u>uncountable</u> set
  - we can't count, but we can still **measure!**

CONTINUOUS RANDOM VARIABLES

Discrete vs. Continuous

Discrete R.V.	Continuous R.V.
Number of heads in n coin tosses	A number from the interval [a,b] where $a,b \in R$
Year of birth of all students in this class	Exact weight of all students in this class
Number of phone calls per minute at a telephone exchange	Time between successive phone calls at a telephone exchange
Winning time of Olympic 100m races rounded to the nearest 100^{th} of a second.	Exact winning time of Olympic 100m races

Mass vs. Density

Any function f(x) where $x \in \mathbb{R}$ with the following two properties:

1) $f(x) \geq 0$
2) $\int f(x)dx = 1$

is called a **probability density** function.

Since its total area is 1, we can treat any area under it as a probability. For example,

$$p(a < X < b) = \int_a^b f(x)\,dx$$

Integration of density implies that density is multiplied by volume (area). Therefore, the resulting probability can be treated as **probability mass** as well.

For continuous random variables, probability density ≠ probability.

Since volume (area) of any point is 0, probability of a single value is 0.

Using calculus as well

$$p(X = a) = \int_a^a f(x)\,dx = 0$$

Mass vs. Density

For discrete random variables, we can use volume of a single point as 1.

Therefore, **probability density = probability mass for discrete random variables**.

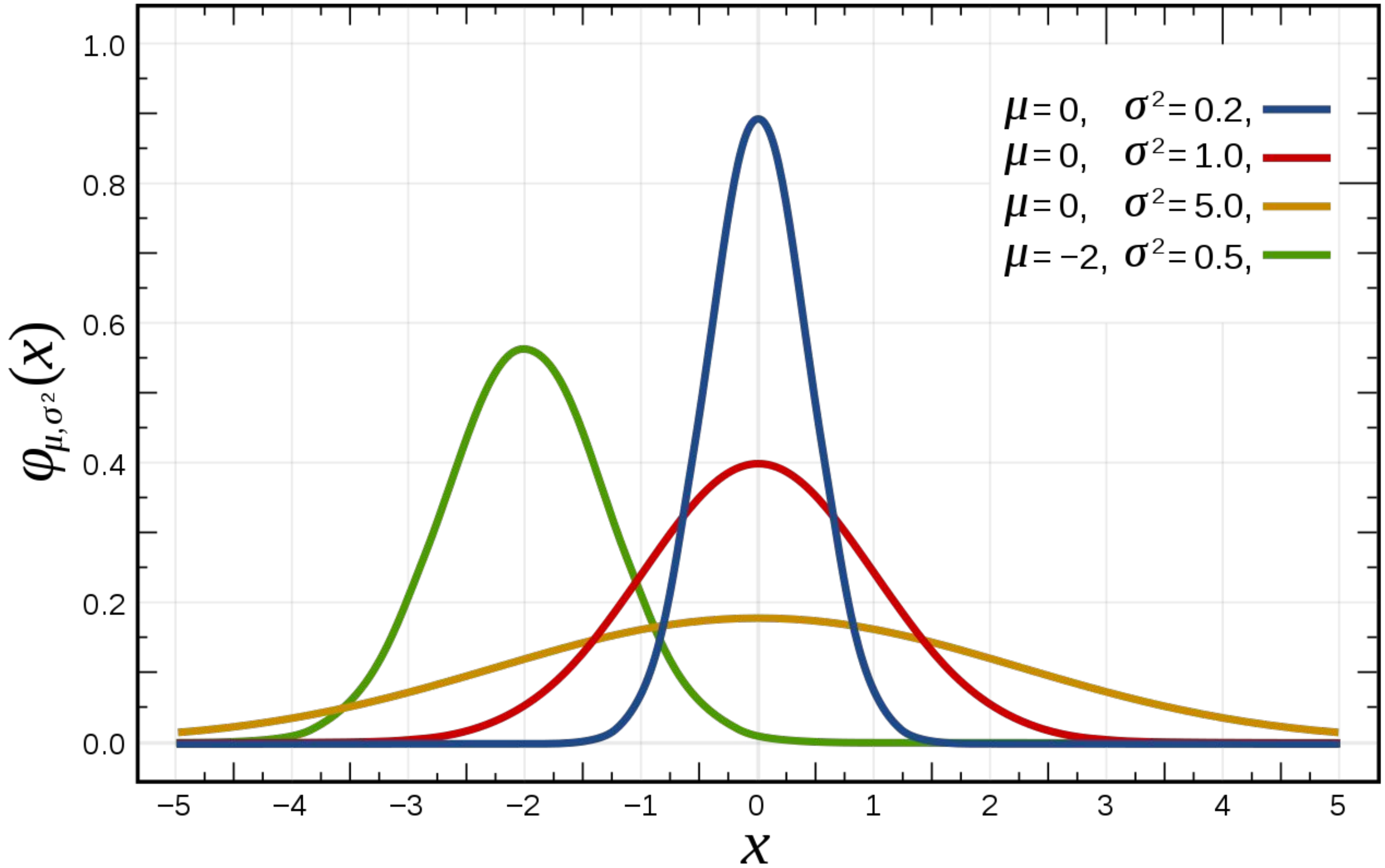
Some books use probability density for discrete random variables.

Remember that in the discrete case, density = probability.

But in the continuous case, density $\neq$ probability.

Normal Density

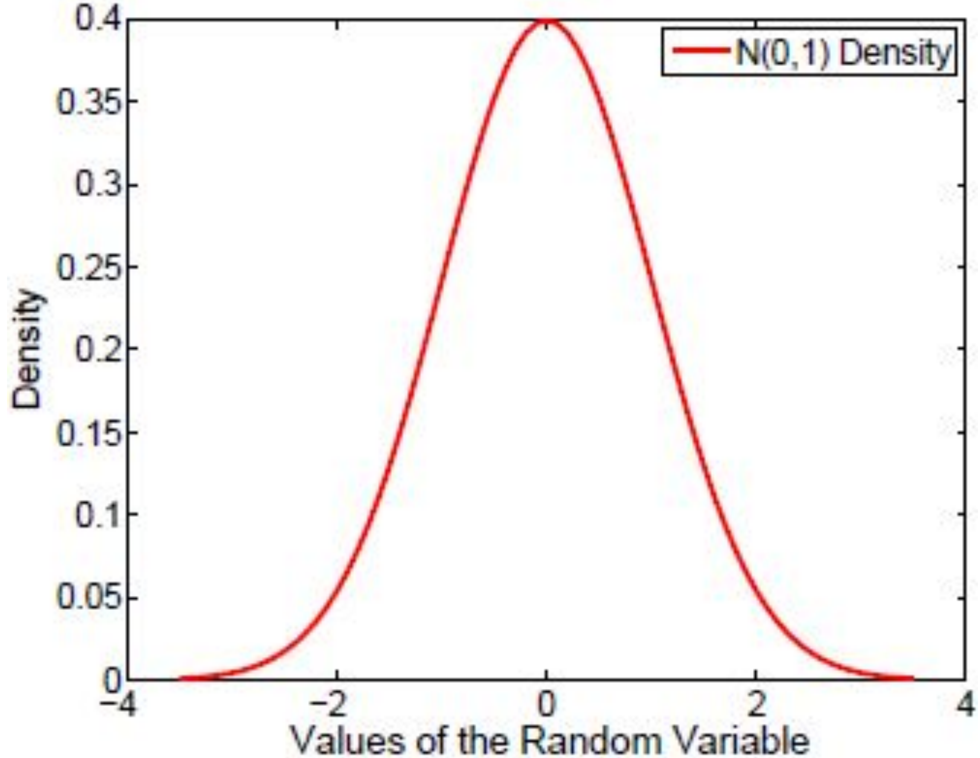
$$N\left(x;\mu,\sigma^2\right) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$



Standard Normal Density

N(x ; 0,1) is the normal density with mean 0 and standard deviation 1.

Known as the **standard normal density**.



Standardization

$$\text{z-score} = (x-\mu)/\sigma$$

Tells you how many standard deviations is x away from the mean.

For example, if $\mu=15$, $\sigma=2$, then $x=11$ has a

z-score = (11-15)/2

= -4/2

= -2 standard deviations away from the mean.

The transformation $Z=(X-\mu)/\sigma$ is called **standardization**.

If random variable X has mean μ and standard deviation σ, then random variable Z will have mean 0 and standard deviation 1.

Standard Normal Table

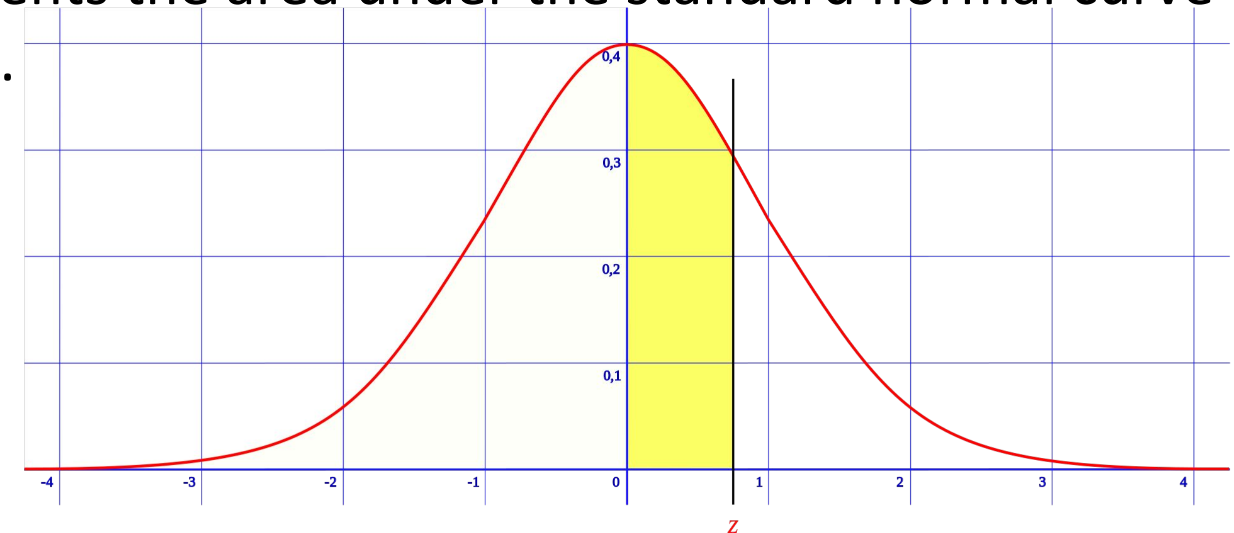
Integral of normal density does not have a simple closed-form formula.

It must be computed numerically.

Fortunately, such computations are already stored in **normal tables**.

The next 4 slides show the **standard normal table**.

Each entry represents the area under the standard normal curve N(0,1) from 0 to z.



z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0	0	0.00399	0.00798	0.01197	0.01595	0.01994	0.02392	0.0279	0.03188	0.03586
0.1	0.03983	0.0438	0.04776	0.05172	0.05567	0.05962	0.06356	0.06749	0.07142	0.07535
0.2	0.07926	0.08317	0.08706	0.09095	0.09483	0.09871	0.10257	0.10642	0.11026	0.11409
0.3	0.11791	0.12172	0.12552	0.1293	0.13307	0.13683	0.14058	0.14431	0.14803	0.15173
0.4	0.15542	0.1591	0.16276	0.1664	0.17003	0.17364	0.17724	0.18082	0.18439	0.18793
0.5	0.19146	0.19497	0.19847	0.20194	0.2054	0.20884	0.21226	0.21566	0.21904	0.2224
0.6	0.22575	0.22907	0.23237	0.23565	0.23891	0.24215	0.24537	0.24857	0.25175	0.2549
0.7	0.25804	0.26115	0.26424	0.2673	0.27035	0.27337	0.27637	0.27935	0.2823	0.28524
0.8	0.28814	0.29103	0.29389	0.29673	0.29955	0.30234	0.30511	0.30785	0.31057	0.31327
0.9	0.31594	0.31859	0.32121	0.32381	0.32639	0.32894	0.33147	0.33398	0.33646	0.33891

z = 0.7+0.02 = 0.72
P(0<x<0.72) = 0.26424

z = 0.9+0.06 = 0.96
P(0<x<0.96) = 0.33147

z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
1	0.34134	0.34375	0.34614	0.34849	0.35083	0.35314	0.35543	0.35769	0.35993	0.36214
1.1	0.36433	0.3665	0.36864	0.37076	0.37286	0.37493	0.37698	0.379	0.381	0.38298
1.2	0.38493	0.38686	0.38877	0.39065	0.39251	0.39435	0.39617	0.39796	0.39973	0.40147
1.3	0.4032	0.4049	0.40658	0.40824	0.40988	0.41149	0.41308	0.41466	0.41621	0.41774
1.4	0.41924	0.42073	0.4222	0.42364	0.42507	0.42647	0.42785	0.42922	0.43056	0.43189
1.5	0.43319	0.43448	0.43574	0.43699	0.43822	0.43943	0.44062	0.44179	0.44295	0.44408
1.6	0.4452	0.4463	0.44738	0.44845	0.4495	0.45053	0.45154	0.45254	0.45352	0.45449
1.7	0.45543	0.45637	0.45728	0.45818	0.45907	0.45994	0.4608	0.46164	0.46246	0.46327
1.8	0.46407	0.46485	0.46562	0.46638	0.46712	0.46784	0.46856	0.46926	0.46995	0.47062
1.9	0.47128	0.47193	0.47257	0.4732	0.47381	0.47441	0.475	0.47558	0.47615	0.4767

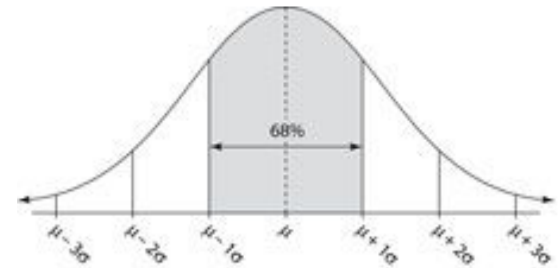
z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
2	0.47725	0.47778	0.47831	0.47882	0.47932	0.47982	0.4803	0.48077	0.48124	0.48169
2.1	0.48214	0.48257	0.483	0.48341	0.48382	0.48422	0.48461	0.485	0.48537	0.48574
2.2	0.4861	0.48645	0.48679	0.48713	0.48745	0.48778	0.48809	0.4884	0.4887	0.48899
2.3	0.48928	0.48956	0.48983	0.4901	0.49036	0.49061	0.49086	0.49111	0.49134	0.49158
2.4	0.4918	0.49202	0.49224	0.49245	0.49266	0.49286	0.49305	0.49324	0.49343	0.49361
2.5	0.49379	0.49396	0.49413	0.4943	0.49446	0.49461	0.49477	0.49492	0.49506	0.4952
2.6	0.49534	0.49547	0.4956	0.49573	0.49585	0.49598	0.49609	0.49621	0.49632	0.49643
2.7	0.49653	0.49664	0.49674	0.49683	0.49693	0.49702	0.49711	0.4972	0.49728	0.49736
2.8	0.49744	0.49752	0.4976	0.49767	0.49774	0.49781	0.49788	0.49795	0.49801	0.49807
2.9	0.49813	0.49819	0.49825	0.49831	0.49836	0.49841	0.49846	0.49851	0.49856	0.49861

z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
3	0.49865	0.49869	0.49874	0.49878	0.49882	0.49886	0.49889	0.49893	0.49896	0.499
3.1	0.49903	0.49906	0.4991	0.49913	0.49916	0.49918	0.49921	0.49924	0.49926	0.49929
3.2	0.49931	0.49934	0.49936	0.49938	0.4994	0.49942	0.49944	0.49946	0.49948	0.4995
3.3	0.49952	0.49953	0.49955	0.49957	0.49958	0.4996	0.49961	0.49962	0.49964	0.49965
3.4	0.49966	0.49968	0.49969	0.4997	0.49971	0.49972	0.49973	0.49974	0.49975	0.49976
3.5	0.49977	0.49978	0.49978	0.49979	0.4998	0.49981	0.49981	0.49982	0.49983	0.49983
3.6	0.49984	0.49985	0.49985	0.49986	0.49986	0.49987	0.49987	0.49988	0.49988	0.49989
3.7	0.49989	0.4999	0.4999	0.4999	0.49991	0.49991	0.49992	0.49992	0.49992	0.49992
3.8	0.49993	0.49993	0.49993	0.49994	0.49994	0.49994	0.49994	0.49995	0.49995	0.49995
3.9	0.49995	0.49995	0.49996	0.49996	0.49996	0.49996	0.49996	0.49996	0.49997	0.49997
4	0.49997	0.49997	0.49997	0.49997	0.49997	0.49997	0.49998	0.49998	0.49998	0.49998

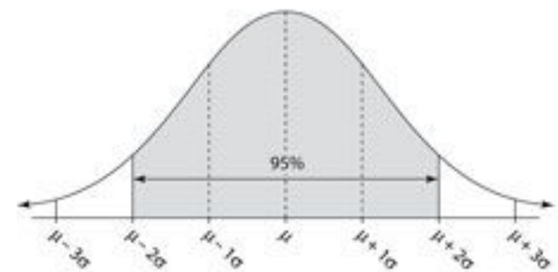
The 68–95–99.7 Rule

In any normal distribution

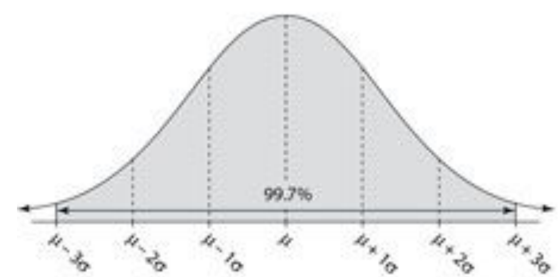
area between ±1 standard deviations from the mean is 68.27%



area between ±2 standard deviations from the mean is 95.45%



area between ±3 standard deviations from the mean is 99.73%



Normal Approximation

Binomial Approximation

The normal distribution can be used as an approximation to the binomial distribution, under certain circumstances, namely:

- If $X \sim B(n, p)$ and if n is large and/or p is close to ½, then X is approximately $N(np, np(1-p))$

In some cases, working out a problem using the Normal distribution may be easier than using a Binomial.

Poisson Approximation

The normal distribution can also be used to approximate the Poisson distribution for large values of λ (the mean of the Poisson distribution).

- If $X \sim Po(\lambda)$ then for large values of λ, $X \sim N(\lambda, \lambda)$ approximately.

Binomial

1) What is the probability of 45 to 50 heads in 100 tosses of a coin with p(H)=0.6?

P(45≤X≤50) =

C(100,45) * .6^45 * .4^55 +

C(100,46) * .6^46 * .4^54 +

C(100,47) * .6^47 * .4^53 +

C(100,48) * .6^48 * .4^52 +

C(100,49) * .6^49 * .4^51 +

C(100,50) * .6^50 * .4^50

2) P(45≤X≤60) will be even more tedious.

3) What is the probability of 6050 to 6070 heads in 10000 tosses of a coin with p(H)=0.6?

C(10000,6050) is not even possible on your calculator.

Normal Approximation

In reality X~Bin(100,0.6)

But using Normal approximation,

X~N(100*0.6, 100*0.6*0.4) = N(60,24)

P(45≤X≤50) = area under N(60,24) curve from 45 to 50.

But since X is actually discrete, we apply a correction for using a continuous approximation.

Since values ≥44.5 are rounded to 45 and values ≤50.5 are rounded to 50, we apply the following **continuity correction**

1) P(44.5 ≤ X ≤ 50.5)

= P((44.5-60)/sqrt(24) ≤ Z ≤ (50.5-60)/sqrt(24))

= P(-3.2639 ≤ Z ≤ -1.9392)

= P(1.9392 ≤ Z ≤ 3.2639)

= P(0 ≤ Z ≤ 3.2639) - P(0 ≤ Z ≤ 1.9392)

= 0.49944 - 0.47381 = 0.02563

2) P(44.5 ≤ X ≤ 60.5) = P(-3.2639 ≤ Z ≤ 0.1021)

= P(0 ≤ Z ≤ 3.2639) + P(0 ≤ Z ≤ 0.1021)

= 0.49944 + 0.03983 = 0.53927

3) P(6049.5 ≤ Y ≤ 6070.5) = P(1.0104 ≤ Z ≤ 1.4391)

= P(0 ≤ Z ≤ 1.4391) - P(0 ≤ Z ≤ 1.0104)

= 0.42507 - 0.34375 = 0.08132

Working with Normal approximation is much easier.
But do not forget the continuity correction.

SOME SPECIAL CONTINUOUS RANDOM VARIABLES

Uniform Random Variable

- Such a random variable takes values in a bounded interval, say (a, b), with density

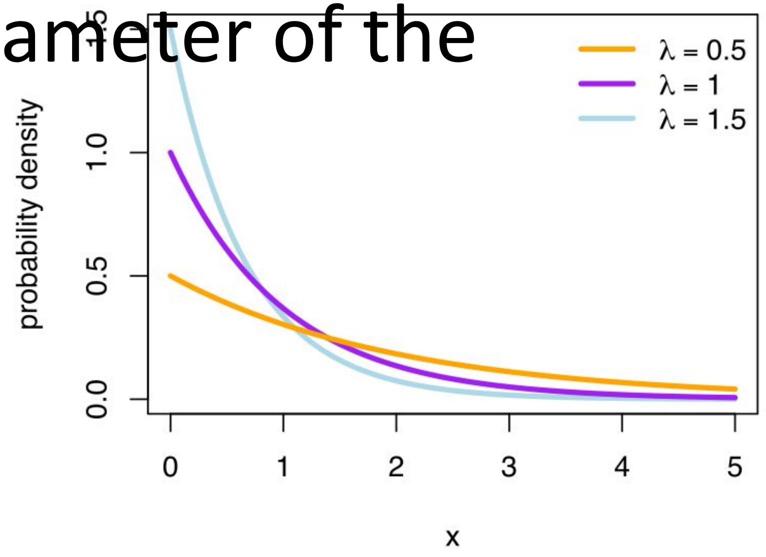
$$f(x) = \begin{cases} \dfrac{1}{b-a}, & \text{for } x \in (a,b) \\ 0, & \text{otherwise} \end{cases}$$

- Denoted by X ~ Uniform(a, b).
- Whenever we say "pick a point randomly …", then the picked point X is a uniform random variable.

Exponential Random Variable

- Takes values in the interval [0, ∞), with density $$f(x)=\begin{cases}\lambda e^{-\lambda x}, & \text{for } x\in(0,\infty)\\ 0, & \text{otherwise}\end{cases}$$

- The constant λ > 0 is a parameter of the density.
- Denoted by X ~ Exp(λ).



Exponential Random Variable

- Time it takes for a particular window glass to crack (due to some accident).
- Time it takes for a bulb to stop working.
- Time it takes for an electrical circuit to malfunction.
- Time it takes for a radioactive atom to decay.

Time between events

Application of Exponential Density – Carbon Dating

- The half-life of the unstable Carbon-14 isotope is roughly around 5730 years.
  - if x amount of carbon-14 material is left to decay naturally, after 5730 years x/2 amount will be left.
- Actual life of a Carbon-14 atom is a random variable with **exponential density**

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

where λ is the rate-of-decay given by $\lambda = \ln(2)/\text{half life} = \ln(2)/5730 = 0.00012097$.

Distribution Function

Let f(x) be **any** density function (non-negative, total area 1).

Let's define a new function

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

F(t) represents the cumulative probability P(X≤t) for $t \in \mathbb{R}$.

Known as the **cumulative distribution function** or just **distribution function**.

Notice that by definition $\frac{d}{dt}F(t) = f(t)$

Properties Of A Distribution Function

- A distribution function, F, always has the following properties
  1. F(t) is a non-decreasing function of t,
  2. $F(-\infty) = 0$, $F(\infty) = 1$,
  3. F(t) is right continuous for all $t \in \mathbb{R}$.
- F(t) is the <u>distribution of a continuous random variable if, **in addition**</u>, there exists a density f, so that $\frac{d}{dt}F(t) = f(t)$ for $t \in \mathbb{R}$.

Continuous R.V Properties

- Range of continuous R.V is an uncountable set.
- <u>Distribution</u> must obey the **fundamental theorem of calculus**

$$P(s < X \le t) = F(t) - F(s) = \int_{s}^{t} f(x)\,dx$$

- For any real number *a*,

$$P(X = a) = P(a < X \le a) = F(a) - F(a) = 0$$

  – Let X be a real number chosen randomly between 5 and 10. Find (i) P(6<X<7)? (ii) P(X=7)?

Deriving a Density

- Suppose we pick a point at random from the interval [3, 14].
  - Sample space is S = [3, 14].
  - X(s)=s   i.e., random variable X is the selected point from [3,14]
- Any subinterval of the form X<=t will have length t-3
- For any subinterval A, P(A) = length(A)/length(S) = length(A)/(14-3)
- Therefore, **distribution function F(t)** of random variable X is

$$F(t) \;=\; \mathbb{P}(X \le t) \;=\; \begin{cases} 0 & \text{if } t < 3, \\ \frac{t-3}{14-3} & \text{if } t \in [3,14], \\ 1 & \text{if } t > 14. \end{cases}$$

- By differentiating the distribution function F(t), we get the **density function** of X

$$f(x) = \begin{cases} \frac{1}{14-3} & \text{if } x \in (3,14), \\ 0 & \text{otherwise.} \end{cases}$$

Example

- Recall that a density function
  - is non-negative
  - with total area 1
- Area from 0 to infinity under $e^{-.01x}$ is given by $\int_0^\infty e^{-.01x}dx = \frac{1}{.01} = 100$
- We can define a **density function** *f(x)* using this result

$$f(x) = \left\{ \begin{matrix} .01e^{-.01x} & ,x \geq 0 \\ 0 & ,x < 0 \end{matrix} \right\}$$

- The **distribution function** F(t)=

$$F(t) = \left\{ \begin{matrix} \int_0^t .01e^{-.01x}dx = \left.\frac{e^{-.01x}}{-.01}\right|_0^t = 1 - e^{-.01t} & ,t \geq 0 \\ 0 & ,t < 0 \end{matrix} \right\}$$

Lifespan of Car Windshields

- It has been empirically observed that the time X it takes for a windshield to develop a crack has density $f(x)=\begin{cases} .01e^{-.01x} & ,x\geq 0 \\ 0 & ,x<0 \end{cases}$ where X is measured in years.
- The probability that a new windshield will crack within *t* years is $P(X\leq t)=F(t)=1-e^{-.01t}$
- Therefore

$$P(X\leq 6\text{ months})=F(.5)=1-e^{-.01(.5)}=0.00499$$
$$P(X\leq 100\text{ years})=F(100)=1-e^{-.01(.5)}=0.63212$$

Other Continuous Random Variables

- Gamma
- Chi-square
- Beta
- Cauchy
- Lognormal
- Logistic