Lecture 13 Normal (Gaussian) Distribution

BIO210 Biostatistics

Xi Chen

Spring, 2023

School of Life Sciences
Southern University of Science and Technology

南方科技大学生命科学学院
SUSTech · SCHOOL OF
LIFE SCIENCES

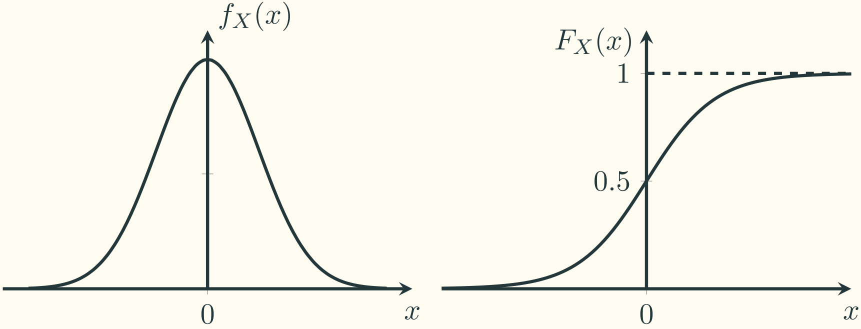
The Normal (Gaussian) PDF

The PDF of a normal distribution

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\dfrac{(x-\mu)^2}{2\sigma^2}}, \qquad \mathbb{E}\left[X\right] = \mu,\ \mathbb{V}\mathrm{ar}\left(X\right) = \sigma^2$$

The Standard Normal (Gaussian) PDF

Standard Normal Distribution: $\mathcal{N}(0,1)$: $f_X(x) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$



General Normal Distribution: $\mathcal{N}(\mu,\sigma^2)$: $f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

The Normal (Gaussian) PDF

We have the random variable $X \sim \mathcal{N}(\mu, \sigma^2)$. Now consider the following random varialbe:

$$Y = aX + b \text{ , where } a \text{ and } b \text{ are constant}$$

- What distribution does Y follow?
- $\mathbb{E}[Y] = ?$
- $\mathbb{V}\mathrm{ar}(Y) = ?$

$Y \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$

Property: A linear function of a normal r.v. is also a normal r.v.

The Normal (Gaussian) PDF

We have the random variable $X \sim \mathcal{N}(\mu, \sigma^2)$. Now consider the following random variable:

$$Z = \frac{X - \mu}{\sigma}$$

- What distribution does Z follow?
- $\mathbb{E}[Z] = ?$
- $\mathbb{V}\mathrm{ar}(Z) = ?$

$Z \sim \mathcal{N}(0, 1)$

The Normal (Gaussian) PDF

Given that X and Y are two independent normal random variables, and $X \sim \mathcal{N}(\mu_x, \sigma_x^2)$ and $Y \sim \mathcal{N}(\mu_y, \sigma_y^2)$, now consider the new random variable:

$$W = X + Y$$

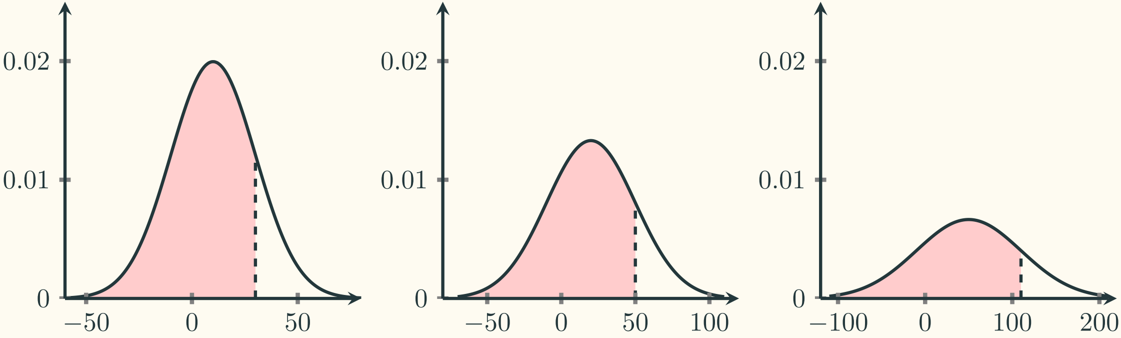
- What distribution does W follow?
- $\mathbb{E}[W] = ?$
- $\mathbb{V}\mathrm{ar}(W) = ?$

$$W \sim \mathcal{N}(\mu_x + \mu_y, \sigma_x^2 + \sigma_y^2)$$

Property: the sum of independent normal random variables is still normal.

Properties of normal PDFs

Dotted line: one standard deviation away from the mean.

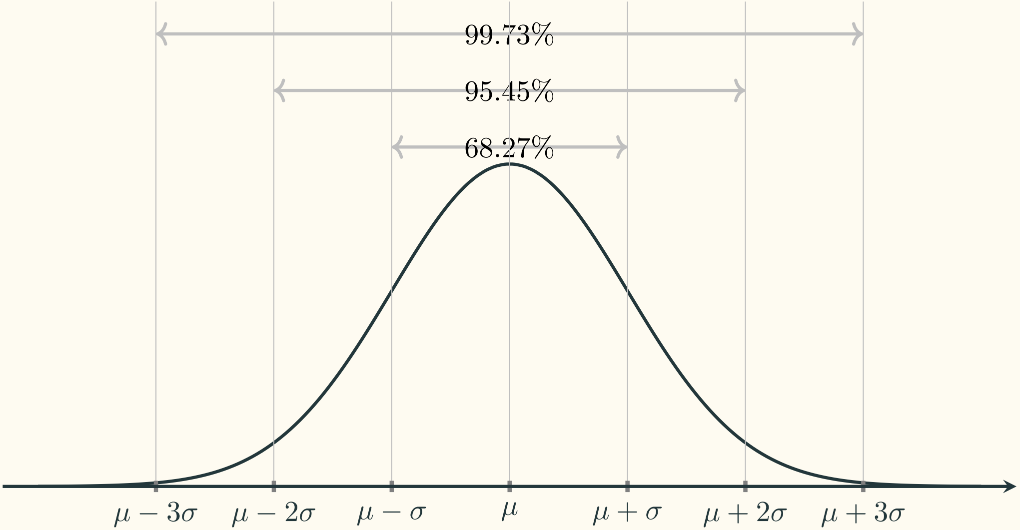


$\mu = 10$ $\sigma = 20$

$\mu = 20$ $\sigma = 30$

$\mu = 50$ $\sigma = 60$

The Empirical Rule



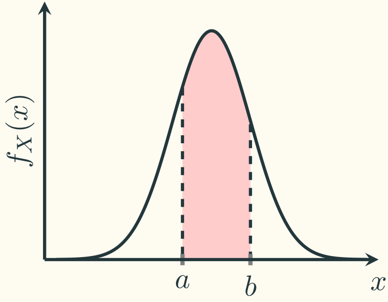
Normal Distribution in real life

- **Commonly observed in many natural phenomena**: height, weight, blood pressure, chest measurements of Scottish soldiers, *etc*.
  - In many cases, you need to take the log value.
- **Noise** or **Error**.
  - An assumption.
- **Sum of many random variables**.
  - Only if they have equal weights.
- **Sample mean**.

TABLE 1: Chest measurement of Scottish soldiers

Girth	Frequency
33	3
34	18
35	81
36	185
37	420
38	749
39	1,073
40	1,079
41	934
42	658
43	370
44	92
45	50
46	21
47	4
48	1

Probability Calculation



$$X \sim \mathcal{N}(\mu, \sigma^2)$$

$$\mathbb{P}\left(a \leqslant X \leqslant b\right) = \int_a^b f_X(x)\,dx$$

$$= \int_a^b \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{-(x-\mu)^2}{2\sigma^2}}\,dx$$

The solution is non-elementary!

Note: we know $\mathbb{P}\left(a \leqslant X \leqslant b\right) = F_X(b) - F_X(a)$

and if $X \sim \mathcal{N}(\mu, \sigma^2)$, then $\frac{X-\mu}{\sigma} \sim \mathcal{N}(0, 1)$.

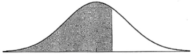
Pre-computed table to the rescue!

Examples of the Standard Normal Table

TABLE A.3
Areas in the upper tail of the standard normal distribution

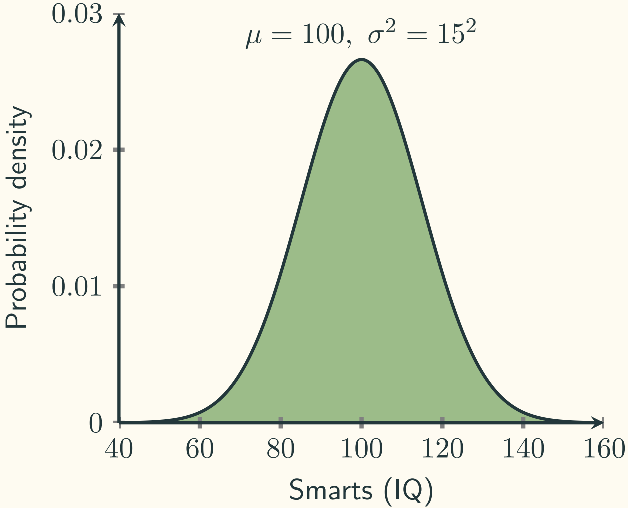
z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.500	0.496	0.492	0.488	0.484	0.480	0.476	0.472	0.468	0.464
0.1	0.460	0.456	0.452	0.448	0.444	0.440	0.436	0.433	0.429	0.425
0.2	0.421	0.417	0.413	0.409	0.405	0.401	0.397	0.394	0.390	0.386
0.3	0.382	0.378	0.374	0.371	0.367	0.363	0.359	0.356	0.352	0.348
0.4	0.345	0.341	0.337	0.334	0.330	0.326	0.323	0.319	0.316	0.312
0.5	0.309	0.305	0.302	0.298	0.295	0.291	0.288	0.284	0.281	0.278
0.6	0.274	0.271	0.268	0.264	0.261	0.258	0.255	0.251	0.248	0.245
0.7	0.242	0.239	0.236	0.233	0.230	0.227	0.224	0.221	0.218	0.215
0.8	0.212	0.209	0.206	0.203	0.200	0.198	0.195	0.192	0.189	0.187
0.9	0.184	0.181	0.179	0.176	0.174	0.171	0.169	0.166	0.164	0.161
1.0	0.159	0.156	0.154	0.152	0.149	0.147	0.145	0.142	0.140	0.138
1.1	0.136	0.133	0.131	0.129	0.127	0.125	0.123	0.121	0.119	0.117
1.2	0.115	0.113	0.111	0.109	0.107	0.106	0.104	0.102	0.100	0.099
1.3	0.097	0.095	0.093	0.092	0.090	0.089	0.087	0.085	0.084	0.082
1.4	0.081	0.079	0.078	0.076	0.075	0.074	0.072	0.071	0.069	0.068
1.5	0.067	0.066	0.064	0.063	0.062	0.061	0.059	0.058	0.057	0.056
1.6	0.055	0.054	0.053	0.052	0.051	0.049	0.048	0.047	0.046	0.046
1.7	0.045	0.044	0.043	0.042	0.041	0.040	0.039	0.038	0.038	0.037
1.8	0.036	0.035	0.034	0.034	0.033	0.032	0.031	0.031	0.030	0.029
1.9	0.029	0.028	0.027	0.027	0.026	0.026	0.025	0.024	0.024	0.023
2.0	0.023	0.022	0.022	0.021	0.021	0.020	0.020	0.019	0.019	0.018
2.1	0.018	0.017	0.017	0.017	0.016	0.016	0.015	0.015	0.015	0.014
2.2	0.014	0.014	0.013	0.013	0.013	0.012	0.012	0.012	0.011	0.011
2.3	0.011	0.010	0.010	0.010	0.010	0.009	0.009	0.009	0.009	0.008
2.4	0.008	0.008	0.008	0.008	0.007	0.007	0.007	0.007	0.007	0.006
2.5	0.006	0.006	0.006	0.006	0.006	0.005	0.005	0.005	0.005	0.005
2.6	0.005	0.005	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004
2.7	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003
2.8	0.003	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002
2.9	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.001	0.001	0.001
3.0	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
3.1	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
3.2	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
3.3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3.4	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

TABLE A.2 Cumulative normal distribution (continued)



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
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
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998
3.5	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998
3.6	.9998	.9998	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999

Example: Human IQ



$$\mathbb{P}(X \geqslant 140) = ?$$

$$\mathbb{P}(130 \leqslant X \leqslant 140) = ?$$

$$Z = \frac{X - \mu}{\sigma} = \frac{140 - 100}{15} = 2.67$$

$$Z = \frac{X - \mu}{\sigma} = \frac{130 - 100}{15} = 2$$

A Historical Fact About The First Standard Normal Table

$$F(x) = \int_0^x e^{-t^2} \mathrm{d}t = x - \frac{x^3}{1!3} + \frac{x^5}{2!5} - \frac{x^7}{3!7} + \frac{x^9}{4!9} - \cdots$$

$$G(x) = \int_x^\infty e^{-t^2} \mathrm{d}t = \frac{1}{x} - \frac{1}{2x^3} + \frac{1 \cdot 3}{4x^5} - \frac{1 \cdot 3 \cdot 5}{8x^7} + \frac{1 \cdot 3 \cdot 5 \cdot 7}{16x^9} - \cdots$$

- Large gaps between $F(x)$ and $G(x)$
- First computed by the French astronomer **Christian Kramp** in 1799.
- Analyse des Réfractions Astronomiques et Terrestres (Analysis of Astronomical and Terrestrial Refractions)

The Table by Christian Kramp

TABLE PREMIÈRE.

Intégrales de $e^{-tt}\,dt$, depuis une valeur quelconque de t jusqu'à t infinie.

t	Intégrale.	Diff. prem.	Diff. II.	Diff. III.
0,00	0,88622692	999968	201	199
0,01	0,87622724	999767	400	199
0,02	0,86622957	999367	599	200
0,03	0,85623590	998768	799	199
0,04	0,84624822	997969	998	197
0,05	0,83626853	996971	1195	199
0,06	0,82629882	995776	1394	196
0,07	0,81634106	994382	1590	195
0,08	0,80639724	992792	1785	194
0,09	0,79646932	991007	1979	195
0,10	0,78655925	989028	2174	192
0,11	0,77666897	986854	2366	190
0,12	0,76680043	984488	2556	189
0,13	0,75695555	981932	2745	188
0,14	0,74713623	979187	2933	186
0,15	0,73734436	976254	3119	184
0,16	0,72758182	973135	3303	183
0,17	0,71785047	969832	3486	180
0,18	0,70815215	966346	3666	175
0,19	0,69848869	962680	3841	178
0,20	0,68886189	958839	4019	173
0,21	0,67927350	954820	4192	171
0,22	0,66972530	950628	4363	168
0,23	0,66021902	946265	4531	166
0,24	0 65075637	941734	4697	163
0,25	0,64133903	937037	4860	160
0,26	0,63196866	932177	5020	157
0,27	0,62264689	927157	5177	155
0,28	0,61337532	929980	5332	151
0,29	0,60415552	916648	5483	149
0,30	0,59498904	911165	5632	145
0,31	0,58587739	905533	5777	142
0,32	0,57682206	899756	5919	138

t	Intégrale.	Diff. prem.	Diff. II.	Diff. III.
0,76	0,25032654	556981	8511	21
0,77	0,24475673	548470	8490	25
0,78	0,23927203	539980	8465	29
0,79	0,23387223	531515	8436	31
0,80	0,22855708	523079	8405	33
0,81	0,22332629	514674	8372	37
0,82	0,21817955	506302	8335	39
0,83	0,21311653	497967	8296	42
0,84	0,20813686	489671	8254	45
0,85	0,20324015	481417	8209	46
0,86	0,19842598	473208	8163	50
0,87	0,19369390	465045	8113	52
0,88	0,18904345	456932	8061	54
0,89	0,18447413	448871	8007	56
0,90	0,17998542	440864	7951	58
0,91	0,17557678	432913	7893	61
0,92	0,17124765	425020	7832	62
0,93	0,16699745	417188	7770	65
0,94	0,16282557	409418	7705	66
0,95	0,15873139	401713	7639	67
0,96	0,15471426	394074	7572	71
0,97	0,15077352	386502	7501	70
0,98	0,14690850	379001	7431	74
0,99	0,14311849	371570	7357	74
1,00	0,13940279	364213	7283	75
1,01	0,13576066	356930	7208	77
1,02	0,13219136	349722	7131	80
1,03	0,12869414	342591	7051	78
1,04	0,12526823	335540	6973	81
1,05	0,12191283	328567	6892	81
1,06	0,11862716	321675	6811	83
1,07	0,11541041	314864	6728	83
1,08	0,11226177	308136	6645	85
1,09	0,10918041	301491	6560	85
1,10	0,10616550	294931	6475	86
1,11	0,10321619	288456	6389	85
1,12	0,10033163	282067	6304	88
1,13	0,09751096	275763	6216	87
1,14	0,09475333	269547	6129	89
1,15	0,09205786	263418	6040	87
1,16	0,08942368	257378	5953	89
1,17	0,08684990	251425	5864	89
1,18	0,08433565	245561	5775	89

t	Intégrale.	Diff. prem.	Diff. II.	Diff. III.	Diff. IV.
2,47	0,00042311518	2186329	105795	4724	191
2,48	0,00040125189	2080534	101071	4533	183
2,49	0,00038044655	1979463	96538	4350	177
2,50	0,00036065192	1882925	92188	4173	171
2,51	0,00034182267	1790737	88015	4002	164
2,52	0,00032391530	1702722	84013	3838	160
2,53	0,00030688808	1618709	80175	3678	152
2,54	0,00029070099	1538534	76497	3526	148
2,55	0,00027531565	1462037	72971	3378	142
2,56	0,00026069528	1389066	69593	3236	137
2,57	0,00024680462	1319473	66357	3099	131
2,58	0,00023360989	1253116	63258	2968	138
2,59	0,00022107873	1189858	60290	2830	141
2,60	0,00020918015	1129568	57460	2749	139
2,61	0,00019788447	1072108	54711	2570	142
2,62	0,00018716339	1017397	52141	2498	118
2,63	0,00017698942	965256	49643	2380	105
2,64	0,00016733686	915613	47263	2275	101
2,65	0,00015818073	868350	44988	2174	95
2,66	0,00014949723	823362	42814	2079	94
2,67	0,00014126361	780548	40735	1985	88
2,68	0,00013345813	739813	38750	1897	85
2,69	0,00012606000	701063	36853	1812	83
2,70	0,00011904937	664210	35041	1729	78
2,71	0,00011240727	629169	33312	1651	76
2,72	0,00010611558	595857	31661	1575	71
2,73	0,00010015701	564196	30086	1504	70
2,74	0,00009451505	534110	28582	1434	67
2,75	0,00008917395	505528	27148	1367	64
2,76	0,00008411867	478380	25781	1303	59
2,77	0,00007933487	452599	24478	1244	60
2,78	0,00007480888	428121	23234	1184	56
2,79	0,00007052767	404887	22050	1128	53
2,80	0,00006647880	382837	20922	1075	51
2,81	0,00006265043	361915	19847	1024	49
2,82	0,00005903128	342068	18823	975	48
2,83	0,00005561060	323245	17848	927	43
2,84	0,00005237815	305397	16921	884	45
2,85	0,00004932418	288476	16037	839	39
2,86	0,00004643942	272439	15198	800	40
2,87	0,00004371503	257241	14398	760	38
2,88	0,00004114262	242843	13638	722	34
2,89	0,00003871419	229205	12916	688	36

Probability Mass/Density Function (PMF/PDF)