LINEAR ALGEBRA

Gram-Schmidt Orthogonalization

Given a basis $\{x_1, x_2, \ldots, x_p\}$ for a subspace W of $\mathbb{R}^n$

$$
\begin{aligned}
v_1 &= x_1 \\
v_2 &= x_2 - \frac{x_2 \cdot v_1}{v_1 \cdot v_1} v_1 \\
v_3 &= x_3 - \frac{x_3 \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{x_3 \cdot v_2}{v_2 \cdot v_2} v_2 \\
&\vdots \\
v_p &= x_p - \frac{x_p \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{x_p \cdot v_2}{v_2 \cdot v_2} v_2 - \cdots - \frac{x_p \cdot v_{p-1}}{v_{p-1} \cdot v_{p-1}} v_{p-1}
\end{aligned}
$$

Then $\{v_1, v_2, \ldots, v_p\}$ is an orthogonal basis for W.

RANDOM VARIABLES

Discrete Random Variables

Let X be a discrete random variable. A function $p(x)$ is a probability mass function of the discrete random variable X if $p(x) \geq 0, \forall x \in X$ and $\sum_x p(x) = 1$.

- Expectation, $E(X) = \sum_x xp(x)$
- If $Y = g(X)$, then $E(Y) = \sum_x g(x)p(x)$
- Variance, $Var(X) = E\left[\left(X - E(X)\right)^2\right] = E(X^2) - [E(X)]^2$
- Standard deviation, $\sigma_X = \sqrt{Var(X)}$
- The cumulative distribution function, $F(t) = P(X \leq t) = \sum_{x \leq t} p(x)$

Continuous Random Variables

Suppose X is a continuous random variable. A function $f(x)$ is called a probability density function of the continuous random variable X if $f(x) \geq 0, \forall x \in X$ and $\int_{-\infty}^{\infty} f(x)dx = 1$.

- Expectation, $E(X) = \int_{-\infty}^{\infty} x\, f(x)dx$
- If $Y = g(X)$, then $E(Y) = \int_{-\infty}^{\infty} g(x)\, f(x)dx$
- Cumulative distribution function, $F(t) = P(X \leq t) = \int_{-\infty}^{t} f(x)dx$

- $P(a \leq X \leq b) = P(a \leq X < b) = P(a < X \leq b) = P(a < X < b) = \int_a^b f(x)dx$

Joint Probability Mass Functions

Suppose X and Y are two discrete random variables. A function $p(x, y)$ is called a joint probability mass function of X and Y if $p(x, y) \geq 0, \forall x \in X, y \in Y$ and $\sum_x \sum_y p(x, y) = 1$.

- Let $Z = g(X, Y)$. Expectation, $E[Z] = \sum_x \sum_y g(x, y)p(x, y)$
- The marginal distributions of X alone and of Y alone are

  $g(x) = \sum_y p(x, y)$ and $h(y) = \sum_x p(x, y)$
- Covariance, $Cov(X, Y) = E(XY) - E(X)E(Y)$
- Correlation of X and Y, $\rho(X, Y) = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$
- If X and Y are independent, then $E(XY) = E(X)E(Y)$

Joint Probability density Functions

Suppose X and Y are two continuous random variables. A function $f(x, y)$ is called a joint probability denisty function of X and Y if $f(x, y) \geq 0, \forall x \in X, y \in Y$ and $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\, dxdy = 1$.

- Let $Z = g(X, Y)$. Expectation, $E[Z] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y)f(x, y)\, dxdy$
- The marginal distributions of X alone and of Y alone are

  $g(x) = \int_{-\infty}^{\infty} f(x, y)\, dy$ and $h(y) = \int_{-\infty}^{\infty} f(x, y)\, dx$
- Covariance, $Cov(X, Y) = E(XY) - E(X)E(Y)$
- Correlation of X and Y, $\rho(X, Y) = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$
- If X and Y are independent, then $E(XY) = E(X)E(Y)$
- The conditional distribution of the random variable Y given that $X = x$ is $f(y|x) = \frac{f(x,y)}{g(x)}$,

  provided $g(x) > 0$
- The conditional distribution of the random variable X given that $Y = y$ is $f(x|y) = \frac{f(x,y)}{h(y)}$,

  provided $h(y) > 0$.

Binomial Distribution

- The probability function of the binomial distribution is given by
  $b(x;n,p) = nC_x\, p^x q^{n-x},\ x = 1,2,3,\cdots$
  Where p is the probability of success and $q = 1 - p$ is the probability of failure.
- Mean, $\mu = np$,
- Variance, $V = \sigma^2 = npq$,
- Standard deviation, $\sigma = \sqrt{npq}$

Poisson Distribution

- The probability function of the Poisson distribution is given by
  $p(x;\lambda) = P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}$, where λ is the parameter of the Poisson distribution.
- Mean, $\mu = \lambda$
- Variance, $V = \sigma^2 = \lambda$
- Standard deviation, $\sigma = \sqrt{\lambda}$

Exponential distribution

- A continuous random variable X assuming non-negative values is said to have an exponential distribution with parameter $\lambda > 0$, if its probability density function is given by

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

- Mean $= \mu = \frac{1}{\lambda}$
- Variance $= \sigma^2 = \left(\frac{1}{\lambda}\right)^2$
- Standard deviation = $\sigma = \frac{1}{\lambda}$

Normal distribution

- A random variable X is said to have a normal distribution with parameters μ (called "mean") and σ^2 (called "variance") if its density function is given by the probability law:

$$n(x;\ \mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}}\ exp\left[\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] \text{ for } -\infty < x < \infty, -\infty < \mu < \infty \text{ and } 0 < \sigma < \infty\,.$$

- Standard normal distribution or $z -$ distribution is given by

$$n(z;\ 0,1) = \frac{1}{\sqrt{2\pi}}\ exp\left[\frac{-z^2}{2}\right], \qquad -\infty < z < \infty,$$

- Cumulative standard normal distribution is

$$\Phi(z) = P(Z \leq z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}}\ e^{-\frac{1}{2}u^2}\ du.$$

$$\Phi(z) = P(Z \le z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2}\, du$$

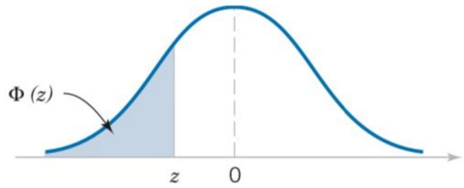

Table III Cumulative Standard Normal Distribution

z	−0.09	−0.08	−0.07	−0.06	−0.05	−0.04	−0.03	−0.02	−0.01	−0.00
−3.9	0.000033	0.000034	0.000036	0.000037	0.000039	0.000041	0.000042	0.000044	0.000046	0.000048
−3.8	0.000050	0.000052	0.000054	0.000057	0.000059	0.000062	0.000064	0.000067	0.000069	0.000072
−3.7	0.000075	0.000078	0.000082	0.000085	0.000088	0.000092	0.000096	0.000100	0.000104	0.000108
−3.6	0.000112	0.000117	0.000121	0.000126	0.000131	0.000136	0.000142	0.000147	0.000153	0.000159
−3.5	0.000165	0.000172	0.000179	0.000185	0.000193	0.000200	0.000208	0.000216	0.000224	0.000233
−3.4	0.000242	0.000251	0.000260	0.000270	0.000280	0.000291	0.000302	0.000313	0.000325	0.000337
−3.3	0.000350	0.000362	0.000376	0.000390	0.000404	0.000419	0.000434	0.000450	0.000467	0.000483
−3.2	0.000501	0.000519	0.000538	0.000557	0.000577	0.000598	0.000619	0.000641	0.000664	0.000687
−3.1	0.000711	0.000736	0.000762	0.000789	0.000816	0.000845	0.000874	0.000904	0.000935	0.000968
−3.0	0.001001	0.001035	0.001070	0.001107	0.001144	0.001183	0.001223	0.001264	0.001306	0.001350
−2.9	0.001395	0.001441	0.001489	0.001538	0.001589	0.001641	0.001695	0.001750	0.001807	0.001866
−2.8	0.001926	0.001988	0.002052	0.002118	0.002186	0.002256	0.002327	0.002401	0.002477	0.002555
−2.7	0.002635	0.002718	0.002803	0.002890	0.002980	0.003072	0.003167	0.003264	0.003364	0.003467
−2.6	0.003573	0.003681	0.003793	0.003907	0.004025	0.004145	0.004269	0.004396	0.004527	0.004661
−2.5	0.004799	0.004940	0.005085	0.005234	0.005386	0.005543	0.005703	0.005868	0.006037	0.006210
−2.4	0.006387	0.006569	0.006756	0.006947	0.007143	0.007344	0.007549	0.007760	0.007976	0.008198
−2.3	0.008424	0.008656	0.008894	0.009137	0.009387	0.009642	0.009903	0.010170	0.010444	0.010724
−2.2	0.011011	0.011304	0.011604	0.011911	0.012224	0.012545	0.012874	0.013209	0.013553	0.013903
−2.1	0.014262	0.014629	0.015003	0.015386	0.015778	0.016177	0.016586	0.017003	0.017429	0.017864
−2.0	0.018309	0.018763	0.019226	0.019699	0.020182	0.020675	0.021178	0.021692	0.022216	0.022750
−1.9	0.023295	0.023852	0.024419	0.024998	0.025588	0.026190	0.026803	0.027429	0.028067	0.028717
−1.8	0.029379	0.030054	0.030742	0.031443	0.032157	0.032884	0.033625	0.034379	0.035148	0.035930
−1.7	0.036727	0.037538	0.038364	0.039204	0.040059	0.040929	0.041815	0.042716	0.043633	0.044565
−1.6	0.045514	0.046479	0.047460	0.048457	0.049471	0.050503	0.051551	0.052616	0.053699	0.054799
−1.5	0.055917	0.057053	0.058208	0.059380	0.060571	0.061780	0.063008	0.064256	0.065522	0.066807
−1.4	0.068112	0.069437	0.070781	0.072145	0.073529	0.074934	0.076359	0.077804	0.079270	0.080757
−1.3	0.082264	0.083793	0.085343	0.086915	0.088508	0.090123	0.091759	0.093418	0.095098	0.096801
−1.2	0.098525	0.100273	0.102042	0.103835	0.105650	0.107488	0.109349	0.111233	0.113140	0.115070
−1.1	0.117023	0.119000	0.121001	0.123024	0.125072	0.127143	0.129238	0.131357	0.133500	0.135666
−1.0	0.137857	0.140071	0.142310	0.144572	0.146859	0.149170	0.151505	0.153864	0.156248	0.158655
−0.9	0.161087	0.163543	0.166023	0.168528	0.171056	0.173609	0.176185	0.178786	0.181411	0.184060
−0.8	0.186733	0.189430	0.192150	0.194894	0.197662	0.200454	0.203269	0.206108	0.208970	0.211855
−0.7	0.214764	0.217695	0.220650	0.223627	0.226627	0.229650	0.232695	0.235762	0.238852	0.241964
−0.6	0.245097	0.248252	0.251429	0.254627	0.257846	0.261086	0.264347	0.267629	0.270931	0.274253
−0.5	0.277595	0.280957	0.284339	0.287740	0.291160	0.294599	0.298056	0.301532	0.305026	0.308538
−0.4	0.312067	0.315614	0.319178	0.322758	0.326355	0.329969	0.333598	0.337243	0.340903	0.344578
−0.3	0.348268	0.351973	0.355691	0.359424	0.363169	0.366928	0.370700	0.374484	0.378281	0.382089
−0.2	0.385908	0.389739	0.393580	0.397432	0.401294	0.405165	0.409046	0.412936	0.416834	0.420740
−0.1	0.424655	0.428576	0.432505	0.436441	0.440382	0.444330	0.448283	0.452242	0.456205	0.460172
0.0	0.464144	0.468119	0.472097	0.476078	0.480061	0.484047	0.488033	0.492022	0.496011	0.500000

$$\Phi(z) = P(Z \le z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} \, du$$

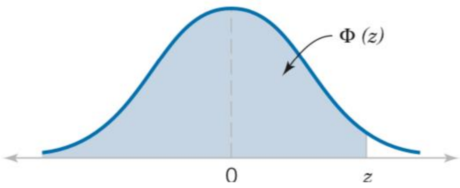

Table III Cumulative Standard Normal Distribution (*continued*)

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.500000	0.503989	0.507978	0.511967	0.515953	0.519939	0.532922	0.527903	0.531881	0.535856
0.1	0.539828	0.543795	0.547758	0.551717	0.555760	0.559618	0.563559	0.567495	0.571424	0.575345
0.2	0.579260	0.583166	0.587064	0.590954	0.594835	0.598706	0.602568	0.606420	0.610261	0.614092
0.3	0.617911	0.621719	0.625516	0.629300	0.633072	0.636831	0.640576	0.644309	0.648027	0.651732
0.4	0.655422	0.659097	0.662757	0.666402	0.670031	0.673645	0.677242	0.680822	0.684386	0.687933
0.5	0.691462	0.694974	0.698468	0.701944	0.705401	0.708840	0.712260	0.715661	0.719043	0.722405
0.6	0.725747	0.729069	0.732371	0.735653	0.738914	0.742154	0.745373	0.748571	0.751748	0.754903
0.7	0.758036	0.761148	0.764238	0.767305	0.770350	0.773373	0.776373	0.779350	0.782305	0.785236
0.8	0.788145	0.791030	0.793892	0.796731	0.799546	0.802338	0.805106	0.807850	0.810570	0.813267
0.9	0.815940	0.818589	0.821214	0.823815	0.826391	0.828944	0.831472	0.833977	0.836457	0.838913
1.0	0.841345	0.843752	0.846136	0.848495	0.850830	0.853141	0.855428	0.857690	0.859929	0.862143
1.1	0.864334	0.866500	0.868643	0.870762	0.872857	0.874928	0.876976	0.878999	0.881000	0.882977
1.2	0.884930	0.886860	0.888767	0.890651	0.892512	0.894350	0.896165	0.897958	0.899727	0.901475
1.3	0.903199	0.904902	0.906582	0.908241	0.909877	0.911492	0.913085	0.914657	0.916207	0.917736
1.4	0.919243	0.920730	0.922196	0.923641	0.925066	0.926471	0.927855	0.929219	0.930563	0.931888
1.5	0.933193	0.934478	0.935744	0.936992	0.938220	0.939429	0.940620	0.941792	0.942947	0.944083
1.6	0.945201	0.946301	0.947384	0.948449	0.949497	0.950529	0.951543	0.952540	0.953521	0.954486
1.7	0.955435	0.956367	0.957284	0.958185	0.959071	0.959941	0.960796	0.961636	0.962462	0.963273
1.8	0.964070	0.964852	0.965621	0.966375	0.967116	0.967843	0.968557	0.969258	0.969946	0.970621
1.9	0.971283	0.971933	0.972571	0.973197	0.973810	0.974412	0.975002	0.975581	0.976148	0.976705
2.0	0.977250	0.977784	0.978308	0.978822	0.979325	0.979818	0.980301	0.980774	0.981237	0.981691
2.1	0.982136	0.982571	0.982997	0.983414	0.983823	0.984222	0.984614	0.984997	0.985371	0.985738
2.2	0.986097	0.986447	0.986791	0.987126	0.987455	0.987776	0.988089	0.988396	0.988696	0.988989
2.3	0.989276	0.989556	0.989830	0.990097	0.990358	0.990613	0.990863	0.991106	0.991344	0.991576
2.4	0.991802	0.992024	0.992240	0.992451	0.992656	0.992857	0.993053	0.993244	0.993431	0.993613
2.5	0.993790	0.993963	0.994132	0.994297	0.994457	0.994614	0.994766	0.994915	0.995060	0.995201
2.6	0.995339	0.995473	0.995604	0.995731	0.995855	0.995975	0.996093	0.996207	0.996319	0.996427
2.7	0.996533	0.996636	0.996736	0.996833	0.996928	0.997020	0.997110	0.997197	0.997282	0.997365
2.8	0.997445	0.997523	0.997599	0.997673	0.997744	0.997814	0.997882	0.997948	0.998012	0.998074
2.9	0.998134	0.998193	0.998250	0.998305	0.998359	0.998411	0.998462	0.998511	0.998559	0.998605
3.0	0.998650	0.998694	0.998736	0.998777	0.998817	0.998856	0.998893	0.998930	0.998965	0.998999
3.1	0.999032	0.999065	0.999096	0.999126	0.999155	0.999184	0.999211	0.999238	0.999264	0.999289
3.2	0.999313	0.999336	0.999359	0.999381	0.999402	0.999423	0.999443	0.999462	0.999481	0.999499
3.3	0.999517	0.999533	0.999550	0.999566	0.999581	0.999596	0.999610	0.999624	0.999638	0.999650
3.4	0.999663	0.999675	0.999687	0.999698	0.999709	0.999720	0.999730	0.999740	0.999749	0.999758
3.5	0.999767	0.999776	0.999784	0.999792	0.999800	0.999807	0.999815	0.999821	0.999828	0.999835
3.6	0.999841	0.999847	0.999853	0.999858	0.999864	0.999869	0.999874	0.999879	0.999883	0.999888
3.7	0.999892	0.999896	0.999900	0.999904	0.999908	0.999912	0.999915	0.999918	0.999922	0.999925
3.8	0.999928	0.999931	0.999933	0.999936	0.999938	0.999941	0.999943	0.999946	0.999948	0.999950
3.9	0.999952	0.999954	0.999956	0.999958	0.999959	0.999961	0.999963	0.999964	0.999966	0.999967

Numerical summation of data

If the n observations in a sample are denoted by $x_1, x_2, \dots, x_n$, the

Sample mean: $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$

Sample variance: $s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$

Standard deviation is the positive square root of the sample variance.

If the population is finite with size N, then for the sampling distribution of $\bar{x}$:

Mean: $\mu_{\bar{x}} = \mu$

Variance: $\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} \times \frac{N-n}{N-1}$, if the sampling is without replacement

$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}$, if the sampling is with replacement or population is infinite.

Sampling distributions

Let X be a random variable of a population with mean μ and variance σ^2.

If the random variables $X_1, X_2, \dots, X_n$ are a random sample on size n,

Sample mean: $\bar{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}$

Sampling distribution of $\bar{X}$

Mean: $\mu_{\bar{X}} = \mu$

Variance: $\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}$

If the population is infinite or sample size is sufficiently large, then the distribution of $Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ is approximately standard normal.

Sampling Distribution of a Difference in Sample Means

Mean: $\mu_{(\bar{X}_1 - \bar{X}_2)} = \mu_{\bar{X}_1} - \mu_{\bar{X}_2} = \mu_1 - \mu_2$

Variance: $\sigma_{(\bar{X}_1 - \bar{X}_2)}^2 = \sigma_{\bar{X}_1}^2 + \sigma_{\bar{X}_2}^2 = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$

If the populations are infinite or sample sizes are sufficiently large, then the distribution of $Z = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$ is approximately standard normal.

Sampling Distribution of Proportions:

Mean $\mu_{\hat{P}} = p$

Standard deviation $\sigma_{\hat{P}} = \sqrt{\frac{pq}{n}} = \left[\frac{p(1-p)}{n}\right]^{\frac{1}{2}}$

The distribution of $Z = \frac{\hat{P}-p}{\sqrt{\frac{p(1-p)}{n}}}$ is approximately standard normal if n is large or $np \geq 5$ and $np(1-p) \geq 5$.