



155+

STATISTICS SYMBOLS

Cheat Sheet

Present by

SymbolALL

Table of Contents

Random variables	6
Values of random variable	6
Sample size	6
Frequency of data	6
Population mean	6
Population standard deviation	6
Sample standard deviation	6
Population proportion	6
Sample proportion	7
Probability of success	7
Probability of failure	7
Population correlation	7
Sample correlation	7
Z-score	7
Significance level (probability of type I error)	7
Probability of type II error	7
Sample regression coefficient	7
Population regression coefficient, Standardized regression coefficient	8
Degree of freedom (df)	8
Sample space	8
Outcome from sample space	8
Population parameters	8
Factorial	9
Double factorial	9
Number of derangements of n objects	9
Permutation (n permute r).....	9
Combination (n choose r).....	9
Multinomial coefficient.....	9

Multiset coefficient (n multichoose r)	9
Probability of event A	10
Complementary probability (Probability of ‘not A ’)	10
Disjunctive probability (Probability of ‘ A or B ’)	10
Joint probability (Probability of ‘ A and B ’)	10
Conditional probability (Probability of ‘ A given B ’)	10
Mean / Expected value of random variable X	10
Conditional expectation (Expected value of X given Y)	11
Variance of random variable X	11
Conditional variance (Variance of X given Y)	11
Standard deviation of random variable X	11
Moment coefficient of skewness of X	11
Kurtosis of random variable X	11
nth central moment of random variable X	11
nth standardized moment of random variable X	11
Covariance of random variables X and Y	12
Correlation of random variables X and Y	12
Probability mass function (pmf) / probability density function (pdf)	13
Support of random variable X	13
Cumulative distribution function (cdf) of random variable X	13
Survival function of random variable X	13
Joint probability function of random variables X_1, \dots, X_n	13
Joint cumulative distribution function of random variables X_1, \dots, X_n	13
Moment-generating function of random variable X	14
Characteristic function of random variable X	14
Cumulant-generating function of random variable X	14
Likelihood function of random variable X with parameter θ	14
Discrete uniform distribution from a to b	15
Bernoulli distribution with p probability of success	15
Geometric distribution with p probability of success	15

Binomial distribution with n trials and p probability of success	15
Negative binomial distribution with r successes and p probability of success	15
Poisson distribution with rate λ	16
Hypergeometric distribution with n draws and K favorable items among N ..	16
Continuous uniform distribution from a to b	18
Exponential distribution with rate λ	18
Normal distribution with mean μ and standard deviation σ	18
Standard normal distribution	18
Pdf of standard normal distribution	18
Cdf of standard normal distribution	18
Error function	18
Positive Z-score associated with significance level α	19
Lognormal distribution with parameters μ and σ	19
Cauchy distribution with parameters x_0 and γ	19
Beta distribution with parameters α and β	19
Beta function	19
Gamma distribution with parameters α and β	19
Gamma function	19
T-distribution with degree of freedom ν	19
Positive t-score with significance level α and degree of freedom ν	20
Chi-squared distribution with degree of freedom ν	20
Chi-squared score with significance level α and degree of freedom ν	20
F-distribution with degrees of freedom ν_1 and ν_2	20
F-score with significance level α and degrees of freedom ν_1 and ν_2	20
I-th value of data set X	21
Sample mean of data set X	21
Median of data set X	21
I-th quartile	21
I-th percentile	21
Sample standard deviation of i-th sample.....	21

Population standard deviation of i-th sample	21
Sample variance	21
Pooled sample variance	22
Population variance	22
Coefficient of determination	22
Eta-squared (Measure of effect size)	22
Predicted average value of y in regression	22
Residual in regression	22
Estimator of parameter θ	22
Bias of estimator $\hat{\theta}$ with respect to parameter θ.....	22
K-th order statistics	22
Events A and B are independent	23
Conditional independence (A and B are independent given C).....	23
Event A increases the likelihood of event B	23
Event A decreases the likelihood of event B	23
Random variable X follows probability distribution F	23
Interquartile range	24
Standard deviation	24
Coefficient of variation	24
Standard error	24
Sum of squares	24
Mean square error	24
Odds ratio	24
Null hypothesis	24
Alternative hypothesis	25
Confidence interval	25
Prediction interval	25
Random variable	25
Independent and identically distributed random variables	25
Law of large numbers	25

List of Probability and Statistics Symbols

Symbol Name	Used For	Example
X, Y, Z, T	Random variables	$E(X_1+X_2)=E(X_1)+E(X_2)$
x, y, z, t	Values of random variable	For all $x \in N(0, P(X=x) = (0.25)x(0.75)$.
n	Sample size	$X-n=X_1+\dots+X_n$
f	Frequency of data	$f_1+\dots+f_k=n$
μ (Mu)	Population mean	$H_0: \mu_1=\mu_2$
σ (Sigma)	Population standard deviation	$\sigma X=\sum(X_i-\mu_x)^2 n$
s	Sample standard deviation	$s=\sqrt{\sum(X_i-\bar{X})^2 n-1}$
π (Pi)	Population proportion	$H_a: \pi_1 \neq \pi_2$

p^{\wedge}	Sample proportion	If $\pi_1=\pi_2$, use $p^{\wedge}=\frac{x_1+x_2}{n_1+n_2}$ instead of p^{\wedge}_1 or p^{\wedge}_2 .
p	Probability of success	In a standard die-tossing experiment, $p=1/6$.
q	Probability of failure	$q=1-p$
ρ (Rho)	Population correlation	$\rho_{X,X}=1$
r	Sample correlation	$r_{xy}=r_{yx}$
z	Z-score	$z=x-\mu\sigma$
α (Alpha)	Significance level (probability of type I error)	At $\alpha=0.05$, the null hypothesis is rejected, but not at $\alpha=0.01$.
β (Beta)	Probability of type II error	$P(H_0 \text{ rejected} H_0 \text{ false}) = 1 - \beta$
b	Sample regression coefficient	$y=b_0+b_1x_1+b_2x_2$

β (Beta)	Population regression coefficient, Standardized regression coefficient	If $\beta_1=0.51$ and $\beta_2=0.8$, then x_2 has more “influence” on y than x_1 .
ν (Nu)	Degree of freedom (df)	$\text{Gamma}(\nu/2, 1/2) = \chi^2(\nu)$
Ω (Capital omega)	Sample space	For a double-coin-toss experiment, $\Omega=\{\text{HH}, \text{HT}, \text{TH}, \text{TT}\}$.
ω (Omega)	Outcome from sample space	$P(X \in A) = P(\{\omega \in \Omega X(\omega) \in A\})$
θ (Theta), β (Beta)	Population parameters	For normal distributions, $\theta=(\mu, \sigma)$.

Operators

Combinatorial Operators

Symbol Name	Explanation	Example
$n!$	Factorial	$4!=4 \cdot 3 \cdot 2 \cdot 1$
$n!!$	Double factorial	$8!!=8 \cdot 6 \cdot 4 \cdot 2$
$!n$	Number of de-range- ments of n objects	Since $\{a,b,c\}$ has 2 permutations where all letter positions are changed, $!3=2$.
nPr	Permutation (n permute r)	$6P3=6 \cdot 5 \cdot 4$
$nCr, (nr)$	Combination (n choose r)	$(nk)=(nn-k)$
(nr_1, \dots, rk)	Multinomial coefficient	$(105,3,2)=10!5!3!2!$
$((nr))$	Multiset coefficient (n multichoose r)	From a 5-element-set, ((53)) 3-element-multisets can be taken.

Probability-related Operators

Symbol Name	Explanation	Example
$P(A)$, $\Pr(A)$	Probability of event A	$P(X \geq 5) = 1 - P(X < 5)$
$P(A')$, $P(A_c)$	Complementary probability (Probability of 'not A ')	For all events E , $P(E) + P(E') = 1$.
$P(A \cup B)$	Disjunctive probability (Probability of 'A or B')	$P(A \cup B) \geq \max(P(A), P(B))$
$P(A \cap B)$	Joint probability (Probability of 'A and B')	Events A and B are mutually exclusive when $P(A \cap B) = 0$.
$P(A B)$	Conditional probability (Probability of 'A given B')	$P(A B) = P(A \cap B) / P(B)$
$E[X]$	Mean / Expected value of random variable X	$E[2f(X) + 5] = 2E[f(X)] + 5$

$E[X Y]$	Conditional expectation (Expected value of X given Y)	$E[X Y=1] \neq E[X Y=2]$
$V(X), \text{Var}(X)$	Variance of random variable X	$V(X) = E[X^2] - E[X]^2$
$V(X Y), \text{Var}(X Y)$	Conditional variance (Variance of X given Y)	$V[X Y] = E[(X - E[X Y])^2 Y]$
$\sigma(X), \text{Std}(X)$	Standard deviation of random variable X	$\sigma(-2X) = -2 \sigma(X)$
Skew[X]	Moment coefficient of skewness of X	$\text{Skew}[X] = E[(X - \mu\sigma)^3]$
Kurt[X]	Kurtosis of random variable X	$\text{Kurt}[X] = E[(X - \mu\sigma)^4]$
$\mu n(X)$	nth central moment of random variable X	$\mu n(X) = E[(X - E[X])n]$
$\mu \sim n(X)$	nth standardized moment of random variable X	$\mu \sim n(X) = E[(X - \mu\sigma)n]$

$\sigma(X,Y)$, $\text{Cov}(X,Y)$	Covariance of random variables X and Y	$\text{Cov}(X,Y) = \text{Cov}(Y,X)$
$\rho(X,Y)$, $\text{Corr}(X,Y)$	Correlation of random variables X and Y	$\rho(X,Y) = \text{Cov}(X,Y)\sigma(X)\sigma(Y)$

Probability-related Functions

Symbol Name	Explanation	Example
$fX(x)$	Probability mass function (pmf) / probability density function (pdf)	$P(Y \leq 2) = \int_{-\infty}^2 fY(y) dy$
RX	Support of random variable X	$RX = \{x \in R \mid fX(x) > 0\}$
$FX(x)$	Cumulative distribution function (cdf) of random variable X	$FX(5) = P(X \leq 5)$
$F-(x), S(x)$	Survival function of random variable X	$S(t) = 1 - F(t)$
$f(x_1, \dots, x_n)$	Joint probability function of random variables X_1, \dots, X_n	$f(1,2) = P(X=1, Y=2)$
$F(x_1, \dots, x_n)$	Joint cumulative distribution function of random variables X_1, \dots, X_n	$F(x,y) = P(X \leq x, Y \leq y)$

$MX(t)$	Moment-generating function of random variable X	$MX(t)=E[etX]$
$\varphi X(t)$	Characteristic function of random variable X	$\varphi X(t)=E[eitX]$
$KX(t)$	Cumulant-generating function of random variable X	$KX(t)=\ln \varphi X(t) = \ln E[eitX]$
$L(\theta x)$	Likelihood function of random variable X with parameter θ	If $X \sim \text{Geo}(p)$, then $L(\theta X=3) = P(X=3 p=\theta)$.

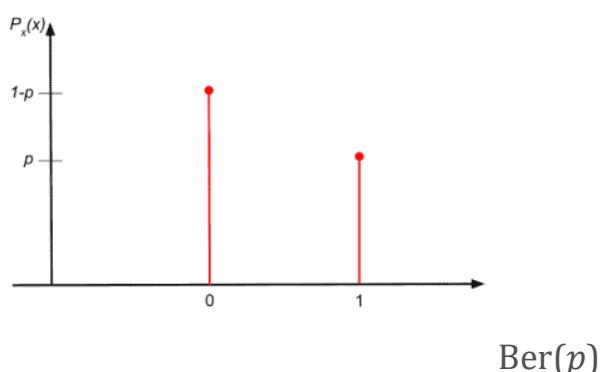
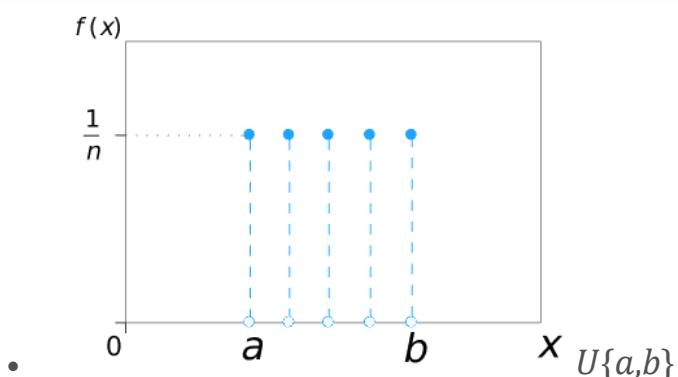
Probability-distribution-related Operators

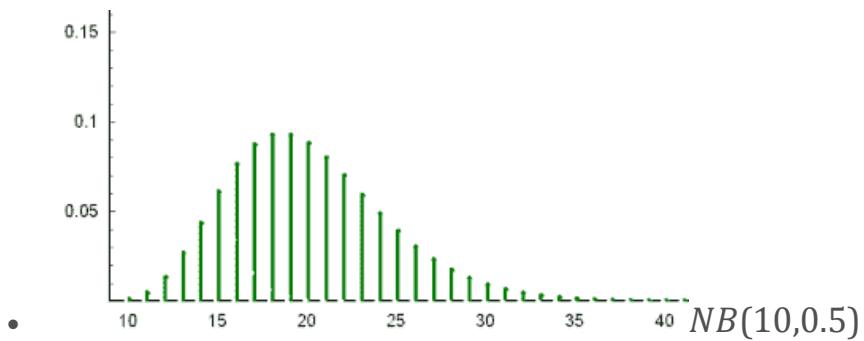
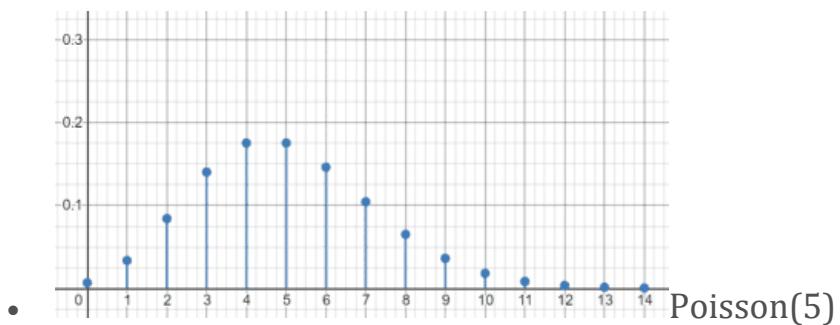
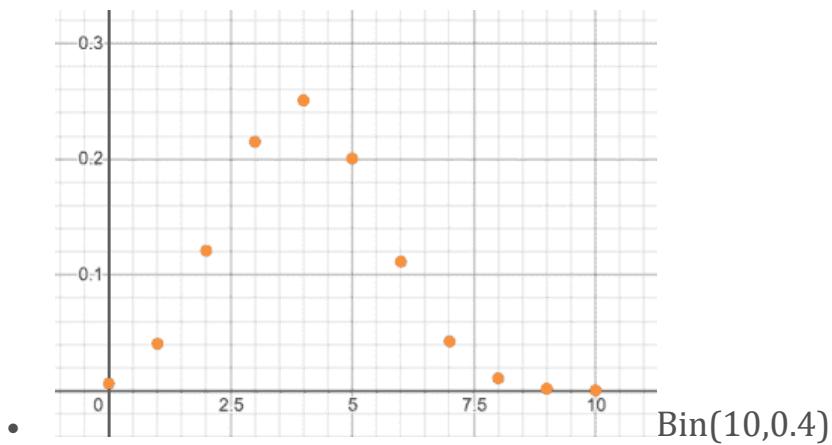
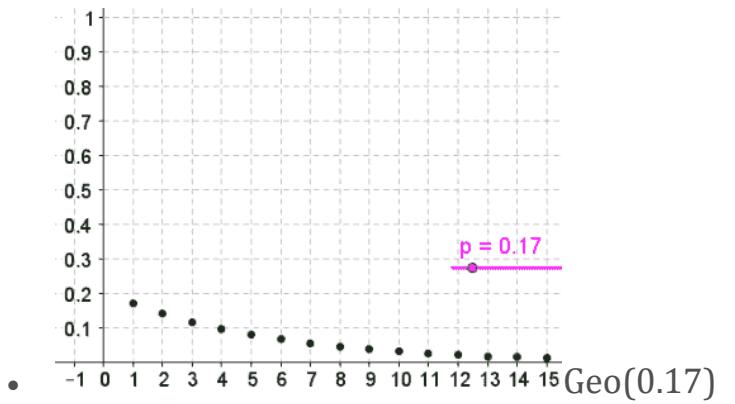
Discrete Probability Distributions

Symbol Name	Explanation	Example
$U\{a,b\}$	Discrete uniform distribution from a to b	Let X be the number on a die following its toss, then $X \sim U\{1,6\}$.
$Ber(p)$	Bernoulli distribution with p probability of success	If $X \sim Ber(0.5)$, then $P(X=0)=P(X=1)=0.5$.
$Geo(p)$	Geometric distribution with p probability of success	If $X \sim Geo(p)$, then $E[X]=1/p$.
$Bin(n,p)$	Binomial distribution with n trials and p probability of success	Let X be the number of heads in a 5-coin toss, then $X \sim Bin(5,0.5)$.
$NB(r,p)$	Negative binomial distribution with r successes and p probability of success	Let Y be the number of die rolls needed to get the third six, then $Y \sim NB(3,1/6)$.

Poisson(λ)	Poisson distribution with rate λ	If $X \sim \text{Poisson}(5)$, then $E[X] = V[X] = 5$.
Hyper(N, K, n)	Hypergeometric distribution with n draws and K favorable items among N	If $X \sim \text{Hyper}(N, K, n)$, then $E[X] = nK/N$.

The following graphs illustrate the **probability mass functions** of 6 of the key distributions mentioned above.





Continuous Probability Distributions and Associated Functions

Symbol Name	Explanation	Example
$U(a,b)$	Continuous uniform distribution from a to b	If $X \sim U(5,15)$, then $P(X \leq 6) = 110$.
$\text{Exp}(\lambda)$	Exponential distribution with rate λ	If $Y \sim \text{Exp}(5)$, then $E[Y] = \sigma[Y] = 15$.
$N(\mu, \sigma^2)$	Normal distribution with mean μ and standard deviation σ	If $X \sim N(1,52)$, then $2X+3 \sim N(5,102)$.
Z	Standard normal distribution	$Z \sim N(0,1)$
$\varphi(x)$	Pdf of standard normal distribution	$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$
$\Phi(x)$	Cdf of standard normal distribution	$\Phi(z) = P(Z \leq z)$
$\text{erf}(x)$	Error function	$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$

$z\alpha$	Positive Z-score associated with significance level α	$z0.025 \approx 1.96$
$\text{Lognormal}(\mu, \sigma^2)$	Lognormal distribution with parameters μ and σ	If $Y \sim \text{Lognormal}(\mu, \sigma^2)$, then $\ln Y \sim N(\mu, \sigma^2)$.
$\text{Cauchy}(x_0, \gamma)$	Cauchy distribution with parameters x_0 and γ	If $X \sim \text{Cauchy}(0, 1)$, then $f(x) = \frac{1}{\pi(x^2+1)}$.
$\text{Beta}(\alpha, \beta)$	Beta distribution with parameters α and β	If $X \sim \text{Beta}(\alpha, \beta)$, then $f(x) \propto x^{\alpha-1}(1-x)^{\beta-1}$.
$B(x, y)$	Beta function	$B(x, y) = \int_0^1 t^{x-1}(1-t)^{y-1} dt$
$\text{Gamma}(\alpha, \beta)$	Gamma distribution with parameters α and β	$\text{Gamma}(1, \lambda) = \text{Exp}(\lambda)$
$\Gamma(x)$	Gamma function	For all $n \in \mathbb{N}_+$, $\Gamma(n) = (n-1)!$
$T(v)$	T-distribution with degree of freedom v	$T(n-1) = X - \bar{X} / S_n$

$t_{\alpha,\nu}$	Positive t-score with significance level α and degree of freedom ν	$t_{0.05,1000} \approx 0.05$
$\chi^2(\nu)$	Chi-squared distribution with degree of freedom ν	$Z_1 + \dots + Z_k = \chi^2(k)$
$\chi_{\alpha,\nu}^2$	Chi-squared score with significance level α and degree of freedom ν	$\chi_{0.05,30}^2 = 43.77$
$F(\nu_1, \nu_2)$	F-distribution with degrees of freedom ν_1 and ν_2	If $X \sim T(\nu)$, then $X^2 \sim F(1, \nu)$.
F_{α,ν_1,ν_2}	F-score with significance level α and degrees of freedom ν_1 and ν_2	$F_{0.05,20,20} \approx 2.1242$

Statistical Operators

Symbol Name	Explanation	Example
X_i, x_i	I-th value of data set X	$x_5=9$
$X-$	Sample mean of data set X	$X-=\sum X_i n$
X^{\sim}	Median of data set X	For a negatively-skewed distribution, $X- \leq X^{\sim}$.
Q_i	I-th quartile	Q_3 is also the 75th (empirical) percentile.
P_i	I-th percentile	$P(X \leq P_{95})=0.95$
s_i	Sample standard deviation of i-th sample	$s_1 > s_2$
σ_i	Population standard deviation of i-th sample	If $\sigma_1=\sigma_2$, then $\sigma_{12}=\sigma_{22}$.
s^2	Sample variance	$s^2=\sum (X_i - X-)^2 n-1$

s^2_p	Pooled sample variance	$s^2_p = \frac{(n_1 - 1)s^2_1 + (n_2 - 1)s^2_2}{n_1 + n_2 - 2}$
σ^2	Population variance	If $\sigma_{12} = \sigma_{22}$, use pooled variance as a better estimate.
r^2, R^2	Coefficient of determination	$R^2 = \frac{SS_{\text{regression}}}{SS_{\text{total}}}$
η^2	Eta-squared (Measure of effect size)	$\eta^2 = \frac{SS_{\text{treatment}}}{SS_{\text{total}}}$
y^{\wedge}	Predicted average value of y in regression	$y^{\wedge} = a + bx_0$
ε^{\wedge}	Residual in regression	$\varepsilon^{\wedge}_i = y_i - y^{\wedge}_i$
θ^{\wedge}	Estimator of parameter θ	If $E(\theta^{\wedge}) = \theta$, then θ^{\wedge} is an unbiased estimator of θ .
$\text{Bias}(\theta^{\wedge}, \theta)$	Bias of estimator θ^{\wedge} with respect to parameter θ	$\text{Bias}(\theta^{\wedge}, \theta) = E[\theta^{\wedge}] - \theta$
$X(k)$	K-th order statistics	$X(n) = \max\{X_1, \dots, X_n\}$

Relational Symbols

Symbol Name	Explanation	Example
$A \perp B$	Events A and B are independent	If $A \perp B$ and $P(A) \neq 0$, then $P(B A) = P(B)$.
$(A \perp B) C$	Conditional independence (A and B are independent given C)	$(A \perp B) C \Leftrightarrow P(A \cap B C) = P(A C)P(B C)$
$A \nearrow B$	Event A increases the likelihood of event B	If $E_1 \nearrow E_2$, then $P(E_2 E_1) \geq P(E_2)$.
$A \searrow B$	Event A decreases the likelihood of event B	If $A \searrow B$, then $A \nearrow B^c$.
$X \sim F$	Random variable X follows probability distribution F	If $X_1, \dots, X_n \sim \text{Ber}(p)$, then $X_1 + \dots + X_n \sim \text{Bin}(n, p)$.

Notational Symbols

Symbol Name	Explanation	Example
IQR	Interquartile range	$IQR=Q_3-Q_1$
SD	Standard deviation	$2SD=2 \cdot 1.5 = 3$
CV	Coefficient of variation	$CV=\sigma/\mu$
SE	Standard error	A statistic of 5.66 corresponds to $10SE$ away from the mean.
SS	Sum of squares	$SS_y = \sum(Y_i - \bar{Y})^2$
MSE	Mean square error	For linear regression, $MSE = \sum(Y_i - \hat{Y}_i)^2 / (n-2)$.
OR	Odds ratio	Let p_1 and p_2 be the rates of accidents in two regions, then $OR = p_1/(1-p_1)p_2/(1-p_2)$.
H_0	Null hypothesis	$H_0: \sigma_{12} = \sigma_{22}$

H_a	Alternative hypothesis	$H_a: \rho > 0$
CI	Confidence interval	$95\%CI = (0.85, 0.97)$
PI	Prediction interval	90%PI is wider than 90%CI, as it predicts an instance of y rather than its average.
r.v.	Random variable	A r. v. is continuous if its support consists of a union of disjoint intervals.
i.i.d.	Independent and identically distributed random variables	If X_1, \dots, X_n are i.i.d. with $V[X_i] = \sigma^2$, then $V[\bar{X}] = \sigma^2/n$.
LLN	Law of large numbers	LLN shows that for all $\varepsilon > 0$, as $n \rightarrow \infty$, $P(\bar{X} - \mu > \varepsilon) \rightarrow 0$.
CLT	Central limit theorem	By CLT, as $n \rightarrow \infty$, $\bar{X} - \mu/\sigma/\sqrt{n} \rightarrow Z$.