MATH 320: Final Study Guide

Lecture 1 – Set Theory (1.1)

How to calculate probability

- Probability by counting equally likely outcomes:

  Probability of an event $= \dfrac{\textit{Number of outcomes in the event}}{\textit{Total number of possible outcomes}}$

- Empirical probability, relative frequency estimate of the probability of an event

  Probability of an event $= \dfrac{\textit{Number of times the event occurs in } n \textit{ trials}}{n}$

Set identities

- Commutative Law (reordering):

  $A \cup B = B \cup A \quad \& \quad A \cap B = B \cap A$

- Associative Law (changing location of parentheses):

  $A \cup (B \cup C) = (A \cup B) \cup C \quad \& \quad A \cap (B \cap C) = (A \cap B) \cap C$

- Distributive Law (distributing union or intersection):

  $A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \quad \& \quad A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

- De Morgan's Law (distributing complement; flip everything):

  $\sim(A \cup B) = \sim A \cap \sim B \quad \& \quad \sim(A \cap B) = \sim A \cup \sim B$

Relationships among sets

- Mutually exclusive (disjoint) if $A \cap B = \emptyset$ (no overlap)
- Pairwise mutually exclusive if $A_i \cap A_j = \emptyset$ for all $i \neq j$ (no overlap of any pairs)
- Exhaustive if $\displaystyle\bigcup_{i=1}^{k} A_i = A_1 \cup \cdots \cup A_k = S$ (complete S)
- Form a partition if exhaustive and pairwise mutually exclusive

Lecture 2 – Counting (1.2)

Basic rules

- Complements counting rule: $n(\sim A) = n(S) - n(A)$
- General union counting rule: $n(A \cup B) = n(A) + n(B) - n(A \cap B)$
- Special case union counting rule: If $A \cap B = \emptyset$, $n(A \cup B) = n(A) + n(B)$
- Union of three events counting rule:

  $n(A \cup B \cup C) = n(A) + n(B) + n(C) - n(A \cap B) - n(A \cap C) - n(B \cap C) + n(A \cap B \cap C)$

Counting principles

- Multiplication principle for counting: If a job consists of k separate tasks, the ith of which can be done in n_i ways ($i = 1, \ldots, k$), then the entire job can be done in $n_1 \times n_2 \times \cdots \times n_k$ ways.
- Ordered with replacement:

  Given n distinguishable objects, there are n^r ways to choose with replacement an ordered sample of r objects.
- Ordered without replacement (all n):

  The number of permutations of n objects is $n! = n(n-1)(n-2)\cdots 2(1)$.
- Ordered without replacement ($r \le n$):

  The number of permutations of n objects taken r at a time is $P\binom{n}{r} = \frac{n!}{(n-r)!}$
- Unordered without replacement ($r \le n$):

  The number of combinations of n objects taken r at a time is $\binom{n}{r} = \frac{n!}{(n-r)!r!}$
- Useful identity (binomial coefficient): $\binom{n}{r} = \binom{n}{n-r}$
- Counting partitions (multinomial coefficient):

  The number of partitions of n objects into k distinct groups of sizes $n_1, n_2, \ldots, n_k$ (where $n_1 + \cdots + n_k = n \iff$ splitting up entire group) is given by: $\binom{n}{n_1,\, n_2,\, \ldots,\, n_k} = \frac{n!}{n_1! n_2! \cdots n_k!}$

<u>**Lecture 3 – Probability**</u> (1.1)

Probability definition based on counting equally likely outcomes

- $P(A) = \dfrac{n(A)}{n(S)}$

Probability when outcomes are not equally likely

- Sample point method:

  Let $S = \{O_1, \ldots, O_n\}$ be a finite set, where all O_i are individual outcomes each with probability $P(O_i) \ge 0$ and $\sum P(O_i) = 1$. For any $A \in S$,
- $P(A) = \displaystyle\sum_{O_i \in A} P(O_i)$

General definition of probability (axioms)

- If you define a way to assign a probability $P(A)$ to any event A, the following axioms must be true
  1. $P(A) \ge 0$
  2. $P(S) = 1$
  3. $P\big(\displaystyle\bigcup_{i=1}^{\infty} A_i\big) = \sum_{i=1}^{\infty} P(A_i)$

Probability theorems

- Complement probability: $P(\sim A) = 1 - P(A)$
- Probability of any event: $P(A) \leq 1$
- Probability of null set: $P(\emptyset) = 0$
- $P(A \cap \sim B) = P(A) - P(A \cap B)$
- General union probability: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- Special case union probability: If $A \cap B = \emptyset$, $P(A \cup B) = P(A) + P(B)$
- Subset probability: If $B \subset A$, then $P(B) \leq P(A)$
- Union of three events probability:

  $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$

Lecture 4 – Conditional Probability (1.3)

Defining conditional probability

- Conditional probability by counting equally likely outcomes: $P(A \mid B) = \dfrac{n(A \cap B)}{n(B)}$
- General definition of conditional probability: $P(A \mid B) = \dfrac{P(A \cap B)}{P(B)}$,

  provided $P(B) > 0$

Probability rules for conditional probability

- All probability theorems hold in conditional probability. Examples below:
- Conditional complement probability: $P(\sim A \mid B) = 1 - P(A \mid B)$
- Conditional general union probability:

  $P(A \cup B \mid C) = P(A \mid C) + P(B \mid C) - P(A \cap B \mid C)$

Multiplication rule for probability

- $P(A \cap B) = P(B)P(A \mid B)$, provided $P(B) > 0$
- $P(A \cap B) = P(A)P(B \mid A)$, provided $P(A) > 0$
- General multiplication rule for probability of k events:

  $P(A_1 \cap \cdots \cap A_k) = P(A_1)P(A_2 \mid A_1) \cdots P(A_k \mid A_1 \cap \cdots \cap A_{k-1})$

Lecture 5 – Independent Events (1.4)

Definition of independence

- Two events A and B, are independent if $P(A \cap B) = P(A)P(B)$

  If $P(A) > 0$ and $P(B) > 0$, then $A \perp\!\!\!\perp B \iff P(A \mid B) = P(A)$, or $P(B \mid A) = P(B)$

  Otherwise, events are said to be dependent. If one condition is true, all are true.

- Special cases of independence:
  - If $P(A) = 0$ or $P(B) = 0$, $A \perp\!\!\!\perp B$
  - If $A \cap B = \emptyset$, $A \perp\!\!\!\perp B$ only if $P(A) = 0$ or $P(B) = 0$
  - If $B \subset A$, $A \perp\!\!\!\perp B$ only if $P(B) = 0$, $P(A) = 0$ or $P(A) = 1$

- Independence of three events: Events A, B, and C are mutually independent if and only if they are pairwise independent (i.e. (A, B), (A, C) and (B, C) are independent pairs) and if $P(A \cap B \cap C) = P(A)P(B)P(C)$.

Applying independence

- Multiplication rule for independent events: If A and B are independent events, $P(A \cap B) = P(A)P(B)$
- Theorems: If A and B are independent events, then the following pairs of events are also independent: A and $\sim B$; $\sim A$ and B; $\sim A$ and $\sim B$

Lecture 6 – Bayes' Theorem (1.5)

Law of total probability

- Let B be an event. If $A_1, \ldots, A_n$ partition the sample space, then

  Law of total probability $=$ P(Second stage event) $= \sum$ Branches of interest

  $$P(B) = P\Big[\bigcup_{i=1}^{n}(A_i \cap B)\Big] = \sum_{i=1}^{n} P(A_i)\,P(B \mid A_i)$$

Bayes' Theorem

- Let B be an event. If $A_1, \ldots, A_n$ partition the sample space, then

  Bayes' Theorem $=$ P(First stage event $\mid$ Second stage event) $= \frac{\text{Main branch of interest}}{\sum \text{All branches of interest}}$

  $$P(A_i \mid B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(A_i)\,P(B \mid A_i)}{\sum_{j=1}^{n} P(A_j)\,P(B \mid A_j)}$$

Lecture 7 – Random Variables (2.1 and 3.1)

Random variables

- Definition: Function from a sample space S into real numbers.
- Range of a RV: The set of possible values of X, $\mathcal{X} = \{x : X(s) = x, s \in S\}$
- RV X is discrete $\iff$ $\mathcal{X}$ is a finite or countable set $\iff$ $F_X(x)$ is a step function of x.
- RV X is continuous $\iff$ $\mathcal{X}$ is an interval (or union of intervals) on the real number line $\iff$ $F_X(x)$ is a continuous function of x.

Lecture 8 – Distribution Functions (2.1 and 3.1)

Calculating probabilities

- Definition: The probability mass function (pmf) of a discrete random variable X is given by

  $f_X(x) = P(X = x), \quad \text{for all } x$
- Definition: A probability density function (pdf) is a continuous random variable X is a real-valued function that can be used to find probabilities using

  $P(a \leq X \leq b) = \int_a^b f(x)\,\mathrm{d}x$

  For $a \in \mathcal{X}$, $P(X = a) = \int_a^a f(x)\,\mathrm{d}x = 0 \quad \Longrightarrow \quad$ For $(a, b) \in \mathcal{X}$, $\quad P(a < X < b) = P(a \leq X \leq b)$

Valid pmfs and pdfs

- Theorem: A function $f_X(x)$ is a pdf (or pmf) of a random variable X if and only if
  (a) $f_X(x) \geq 0$ for all x.
  (b) $\displaystyle\sum_x f_X(x) = 1$ (pmf) or $\int_{-\infty}^{\infty} f_X(x)\,\mathrm{d}x = 1$ (pdf).

Cumulative distribution function (cdf)

- Definition: $F_X(x) = P_X(X \leq x), \quad -\infty < x < \infty$
- Properties of cdfs:
  1. The cdf is defined for $-\infty < x < \infty$ always.
  2. The range of every cdf is $0 \leq F(x) \leq 1 \iff$ Limits: $\displaystyle\lim_{x \to -\infty} F(x) = 0$ and $\displaystyle\lim_{x \to \infty} F(x) = 1$
  3. $F_X(x)$ is a non-decreasing function.
  4. If X is discrete $\rightarrow F(x)$ is a right continuous step function.

     If X is continuous $\rightarrow F(x)$ is a continuous function.
- Relationship between continuous cdf and pdf

  $F'(x) = f(x)$, or equivalently $\frac{\mathrm{d}}{\mathrm{d}x} F_X(x) = f_X(x)$
- Alternate definition of pdf:

  The pdf of a continuous random variable X as the function that satisfies $F_X(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t \quad$ for all x.

Finding probabilities using the cdf

- Cdf always gives a left probability.
- If X is discrete, $F(a) = P(X \leq a) = \sum_{x \leq a} f(x)$

  "Complement of cdf": $1 - F(x) = 1 - P(X \leq x) = P(X > x)$

  Interval probabilities: $P(a < X \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a)$
- If X is continuous: $F_X(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t$

  For a specific value of $x = a$, we find probability with: $F(a) = \int_{-\infty}^{a} f(x)\,\mathrm{d}x$

  Complement of cdf: $1 - F(a) = 1 - P(X \leq a) = 1 - F(a)$

  Interval probabilities: $P(a \leq X \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a)$

<u>**Lecture 9 – Summary Measures**</u> (2.2, 2.3 and 3.1)

Expected value

- Definition:

  If X is discrete $\to \mu = E(X) = \sum x\, f(x)$

  If X is continuous $\to \mu = E(X) = \int_{-\infty}^{\infty} x\, f(x)\,\mathrm{d}x$

Expected value of a function of a random variable

- If $Y = aX + b \to E(Y) = E(aX + b) = aE(X) + b$
- If X is discrete:

  (Used in the derivation of the above identity) If $Y = aX + b \to f_Y(y) = f_Y(ax + b) = f_X(x)$

  In general, if $Y = g(X) \to E(Y) = \sum_y y\, f(y) = E[g(X)] = \sum_x g(x)\, f(x)$
- If X is continuous $\to E(Y) = \int_{-\infty}^{\infty} y\, f(y)\,\mathrm{d}y = E[g(X)] = \int_{-\infty}^{\infty} g(x)\, f(x)\,\mathrm{d}x$
- Linear / Distributive property of expectation:

$$E\left[\sum_{i=1}^{k} c_i g_i(X)\right] = \sum_{i=1}^{k} c_i E[g_i(X)]$$

Variance and standard deviation

- Variance definitions:

$$V(X) = \begin{cases} & \underline{\text{In general}} & & \underline{\text{Discrete}} & \underline{\text{Continuous}} \\ 0) & \sigma^2 & & & \\ 1) & E[(X-\mu)^2] & \to & \sum (x-\mu)^2\, f(x) & \int_{-\infty}^{\infty} (x-\mu)^2\, f(x)\,\mathrm{d}x \\ 2) & E(X^2) - \mu^2 & \to & \sum x^2\, f(x) - \left[\sum x\, f(x)\right]^2 & \int_{-\infty}^{\infty} x^2\, f(x)\,\mathrm{d}x - \left[\int_{-\infty}^{\infty} x\, f(x)\,\mathrm{d}x\right]^2 \end{cases}$$

- Using variance definition 2) $\Rightarrow E(X^2) = V(X) + [E(X)]^2$
- Standard deviation definition: $\sigma = SD(X) = \sqrt{V(X)}$

Variance and standard deviation of $Y = aX + b$

- If $Y = aX + b \rightarrow \sigma_Y^2 = V(Y) = V(aX + b) = a^2 V(X) = a^2 \sigma_X^2$
  $\Rightarrow \sigma_Y = SD(Y) = SD(aX + b) = |a|\, SD(X) = |a|\, \sigma_X$

Mode

- Definition: Mode is the x value which maximizes the distribution function $f(x)$.

Median and Percentiles

- Median m of a continuous random variable X is the solution to: $F(m) = P(X \leq m) = 0.5$.
- Percentile: For $0 \leq p \leq 1$, the $100p^{th}$ percentile of X is the number x_p defined by $F(x_p) = p$.
- $IQR = Q_3 - Q_1$.

Lecture 10 – Discrete Distributions (2.1, 2.3, 2.4, 2.5, 2.6, 2.7)

See distribution table

Summary of four Bernoulli-based experiments

- Distributions: Binomial, Geometric, Negative Binomial, and Hypergeometric
- Throughout all of these, there were three important aspects:
  (1) Number of successes (2) Number of trials (3) Probability of success
- Organization of the four distributions based on what we are interested in (the random variable) and what we are given (as parameters).
  - Distributions counting the number of successes: Binomial and Hypergeometric
    * Interested in: (1)
    * (2) and (3) are given as parameters.
    * Only difference is with vs without replacement
  - Distributions counting the number of trials: Geometric and Negative Binomial
    * Interested in: (2)
    * (1) and (3) are given as parameters.
    * Only difference is the number of successes

Poisson distribution

- The random variable X counts the number of events in a given unit.

Lecture 11 – Continuous Distributions (3.1, 3.2, 3.3)

See distribution table

Survival function

- $S(t) = P(T > t) = 1 - F(t)$.

Linear transformation of normal random variables

- Theorem: If $X \sim \text{Normal}\,(\mu, \sigma^2)$ and $Y = aX + b \rightarrow Y \sim \text{Normal}\,(a\mu + b, a^2\sigma^2)$.
- Standardizing: If $X \sim \text{Normal}\,(\mu, \sigma^2)$ and $Z = \frac{X-\mu}{\sigma} \rightarrow Z \sim \text{Normal}\,(\mu = 0, \sigma^2 = 1)$.
- Can standardize any random variable.

Normal probabilities and percentiles

- Z-table: Gives $F_Z(z) = P(Z \leq z)$.
- $P(x_1 \leq X \leq x_2) = P\left(\frac{x_1-\mu}{\sigma} \leq \frac{X-\mu}{\sigma} \leq \frac{x_2-\mu}{\sigma}\right) = P(z_1 \leq Z \leq z_2)$,

  where $z_1 = \frac{x_1-\mu}{\sigma}$ and $z_2 = \dfrac{x_2-\mu}{\sigma}$.
- $F_Z(z_p) = p \quad \rightarrow \quad z_p = \frac{x_p-\mu}{\sigma} \quad \rightarrow \quad x_p = \sigma z_p + \mu$.

Sums of independent, identically distribution random variables

- Central Limit Theorem (CLT): If $X_i \overset{iid}{\sim} f(x)$ with mean μ and variance σ^2 and $S = \sum_{i=1}^{n} X_i$

  $\rightarrow S \overset{approx}{\sim} \text{Normal}\,(n\mu, n\sigma^2)$ for large n.

Lecture 12 – Moment Generating Functions (2.3, 2.4, 2.6, 2.7, 3.1, 3.2, 3.3)

Moments

- Definition: n^{th} moment of $X = E(X^n), \qquad n = 1, 2, 3\ldots$
- n^{th} moment of X about $b = E[(X-b)^n], \qquad n = 1, 2, 3\ldots$
- Central moments $= E[(X-\mu)^n], \qquad n = 1, 2, 3\ldots$

Moment generating functions (mgf)

- Definition:

$$M_X(t) = \underbrace{E(\mathrm{e}^{tx})}_{\text{In general}} \quad \rightarrow \quad \underbrace{\sum_x \mathrm{e}^{tx} f(x)}_{\text{Discrete}} \quad \underbrace{\int_{-\infty}^{\infty} \mathrm{e}^{tx}\, f(x)\, \mathrm{d}x}_{\text{Continuous}}$$

- How to find moments from mgf:

  $M_X'(0) = E(X), \quad M_X''(0) = E(X^2), \quad \dots \quad , \quad M_X^{(n)}(0) = E(X^n)$

  * If X is discrete $\rightarrow M_X^{(n)}(t) = \sum x^n e^{tx} f(x)$ and $M_X^{(n)}(0) = \sum x^n f(x) = E(X^n)$

- Mgf of $Y = aX + b$

  $M_Y(t) = M_{aX+b}(t) = e^{tb} M_X(at)$

- Mgfs are unique.

- Another variance definition: Using mgfs: $M_X''(0) - \left[M_X'(0)\right]^2 = M_X''(t)\big|_{t=0} - \left[M_X'(t)\big|_{t=0}\right]^2$

Lecture 13 – Functions of Random Variables (5.1)

Expected value of a loss or claim

- In general, if loss x with deductible d and cap c, we have

$$g(x) = \begin{cases} 0 & 0 < x \le d \\ x - d & d < x \le d + c \\ c & x > d + c \end{cases}$$

Distribution functions of transformations

- Theorem: Let X have cdf $F_X(x)$ with range $\mathcal{X}$, $Y = g(X)$ with and range $\mathcal{Y}$ and inverse $h(y)$.

  - If $g(x)$ is strictly increasing on $\mathcal{X} \longrightarrow F_Y(y) = F_X(h(y))$ for $y \in \mathcal{Y}$.
  - If $g(x)$ is strictly decreasing on $\mathcal{X} \longrightarrow F_Y(y) = 1 - F_X(h(y))$ for $y \in \mathcal{Y}$.

  - If $g(x)$ is strictly increasing or strictly decreasing on $\mathcal{X}$, then

$$f_Y(y) = f_X(h(y)) \cdot |h'(y)| \qquad \text{for } y \in \mathcal{Y}.$$

Distributions

Discrete Distributions	
Discrete uniform (N_0, N_1)	
Pmf	$P(X = x \mid N_0, N_1) = \frac{1}{N_1 - N_0 + 1}; \qquad x = N_0, \ldots, N_1; \qquad N_0 \leq N_1$
Mean and Variance	$E(X) = \frac{N_0 + N_1}{2}, \qquad V(X) = \frac{(N_1 - N_0 + 1)^2 - 1}{12}$
Mgf	$M_X(t) = \frac{1}{N_1 - N_0 + 1} \sum_{x=N_0}^{N_1} \mathrm{e}^{tx}$
Notes	
Bernoulli (p)	
Pmf	$P(X = x \mid p) = p^x (1-p)^{1-x}; \qquad x = 0, 1; \qquad 0 < \mathrm{p} < 1$
Mean and Variance	$E(X) = p, \qquad V(X) = p(1-p) = pq$
Mgf	$M_X(t) = (1-p) + p\mathrm{e}^t = q + p\mathrm{e}^t$
Notes	Special case of binomial with $n = 1$.
Binomial (n, p)	
Pmf	$P(X = x \mid n, p) = \binom{n}{x} p^x (1-p)^{n-x}; \qquad x = 0, 1, \ldots, n; \qquad 0 < p < 1$
Mean and Variance	$E(X) = np, \qquad V(X) = np(1-p) = npq$
Mgf	$M_X(t) = (q + p\mathrm{e}^t)^n$
Notes	Sum of *iid* bernoulli RVs.
Geometric (p)	
Pmf	$P(X = x \mid p) = q^{x-1} p; \qquad x = 1, 2, \ldots; \qquad 0 < p < 1$
Cdf	$F_X(x \mid p) = 1 - q^x$
Mean and Variance	$E(X) = \frac{1}{p}, \qquad V(X) = \frac{1-p}{p^2} = \frac{q}{p^2}$
Mgf	$M_X(t) = \frac{p\mathrm{e}^t}{1 - q\mathrm{e}^t}; \qquad t < -\ln(q)$
Notes	Special case of negative binomial with $r = 1$. * See other geometric probabilities. Alternate form $Y = X - 1$. This distribution is *memoryless*: $P(X > s \mid X > t) = P(X > s - t); \qquad s > t$.

Negative binomial (r, p)		
Pmf	$P(X = x \mid r, p) = P(X = x \mid r, p) = \binom{x-1}{r-1} p^r q^{x-r}; \qquad x = r, r+1, \ldots; \qquad 0 < p < 1$	
Mean and Variance	$E(X) = \frac{r}{p}, \qquad V(X) = \frac{r(1-p)}{p^2} = \frac{rq}{p^2}$	
Mgf	$M_X(t) = \left[\frac{pe^t}{1-qe^t}\right]^r; \qquad t < -\ln(q)$	
Notes	Sum of *iid* geometric RVs.	
Hypergeometric (N, M, K)		
Pmf	$P(X = x \mid r, p) = P(X = x \mid N, M, K) = \frac{\binom{M}{x}\binom{N-M}{K-x}}{\binom{N}{K}}; \qquad x = 0, 1, \ldots, \min(M, K)$	
Mean and Variance	$E(X) = K\left(\frac{M}{N}\right), \qquad V(X) = K\left(\frac{M}{N}\right)\left(\frac{N-M}{N}\right)\left(\frac{N-K}{N-1}\right)$	
Mgf		
Notes	If do not require $M \geq K$, $\mathcal{X} = \{\max(0, K + M - N), \ldots, \min(M, K)\}$, mean and variance converge to that of binomial $(n = K, p = M/K)$ when $N \to \infty$.	
Poisson (λ)		
Pmf	$P(X = x \mid \lambda) = \frac{e^{-\lambda}\lambda^x}{x!}, \qquad x = 0, 1, 2, \ldots; \qquad \lambda > 0$	
Mean and Variance	$E(X) = \lambda, \qquad V(X) = \lambda$	
Mgf	$M_X(t) = e^{\lambda(e^t - 1)}$	
Notes	If $X_i \overset{\perp\!\!\!\perp}{\sim}$ Poisson (λ_i), then $\sum X_i \sim \text{Poisson}(\lambda = \sum \lambda_i)$.	

Other geometric probabilities

- Let $X \sim \text{Geometric}(p)$.

$$
\begin{aligned}
P(X < \infty) &= 1 \\
P(X > x) &= q^x \\
P(X \geq x) &= q^{x-1} \\
P(a < X \leq b) &= q^a - q^b \\
P(a \leq X \leq b) &= q^{a-1} - q^b
\end{aligned}
$$

Continuous Distributions

Continuous uniform (a, b)

Pdf	$f(x \mid a,b) = \frac{1}{b-a}, \qquad a \le x \le b; \qquad a,b \in \mathbb{R}, \qquad a \le b$
Cdf	$F(x) = \frac{x-a}{b-a} \qquad a \le x \le b$
Survival	$S(t) = \frac{b-t}{b-a} \qquad a \le t \le b \qquad$ if $T \sim$ Uniform (a, b)
Mean and Variance	$E(X) = \frac{a+b}{2}; \qquad V(X) = \frac{(b-a)^2}{12}$
Mgf	$M_X(t) = \frac{\mathrm{e}^{tb} - \mathrm{e}^{ta}}{t(b-a)} \qquad t \neq 0$
Notes	

Exponential (λ)

Pdf	$f(t \mid \lambda) = \lambda \mathrm{e}^{-\lambda t}, \qquad t \ge 0; \qquad \lambda > 0$
Cdf	$F(t) = 1 - \mathrm{e}^{-\lambda t} \qquad t \ge 0$
Survival	$S(t) = \mathrm{e}^{-\lambda t} \qquad t \ge 0$
Mean and Variance	$E(X) = \frac{1}{\lambda}; \qquad V(X) = \frac{1}{\lambda^2}$
Mgf	$M_X(t) = \frac{\beta}{\beta - t} \qquad t < \beta; \qquad$ if $T \sim$ Exp (β)
Notes	Special case of gamma with $\alpha = 1, \beta$. This distribution is *memoryless*: $P(T > a + b \mid T > a) = P(T > b); \qquad a, b > 0.$ Rate parameterization is given; alternate parameterization is with scale $\theta = 1/\lambda$.

Gamma (α, β)

Pdf	$f(x \mid \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \mathrm{e}^{-\beta x}, \qquad x \ge 0; \qquad \alpha, \beta > 0$
Cdf	N/A
Mean and Variance	$E(X) = \frac{\alpha}{\beta} \qquad V(X) = \frac{\alpha}{\beta^2}$
Mgf	$M_X(t) = \left(\frac{\beta}{\beta - t}\right)^\alpha \qquad t < \beta$
Notes	$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} \mathrm{e}^{-x}\, \mathrm{d}x$ Sum of *iid* exponential RVs. A special case is exponential $(\alpha = 1, \beta)$. Rate parameterization is given; alternate parameterization is with scale $\theta = 1/\beta$.

Normal (μ, σ^2)

Pdf	$f(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], \qquad -\infty < x < \infty; \qquad -\infty < \mu < \infty, \quad \sigma > 0$
Cdf	N/A
Mean and Variance	$E(X) = \mu, \qquad V(X) = \sigma^2$
Mgf	$M_X(t) = \exp\left[\mu t + \frac{\sigma^2 t^2}{2}\right]$
Notes	Special case: Standard normal $Z \sim$ Normal $(\mu = 0, \sigma^2 = 1)$.

Lognormal (μ, σ^2)	
Pdf	$f(y \mid \mu, \sigma^2) = \frac{1}{y\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(\ln(y)-\mu)^2}{2\sigma^2}\right]; \qquad y \geq 0; \qquad -\infty < \mu < \infty; \qquad \sigma > 0$
Mean and Variance	$E(Y) = \mathrm{e}^{\mu+\frac{\sigma^2}{2}}, \qquad V(Y) = \mathrm{e}^{2\mu+\sigma^2}(\mathrm{e}^{\sigma^2} - 1)$
Mgf	
Notes	If $Y \sim$ Lognormal $\Longrightarrow \ln(Y) \sim$ Normal (μ, σ^2); equivalently, if $X \sim$ Normal (μ, σ^2) and $Y = \mathrm{e}^X \Longrightarrow Y \sim$ Lognormal. μ and σ^2 represent the mean and variance of the normal random variable X which appears in the exponent.
Beta (α, β)	
Pdf	$f(x \mid \alpha, \beta) = \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}; \qquad 0 \leq x \leq 1; \qquad \alpha, \beta > 0$
Mean and Variance	$E(X) = \frac{\alpha}{\alpha+\beta}, \qquad V(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$
Mgf	
Notes	$B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\,\mathrm{d}x = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$