

Important Distributions

The nine most important distributions we have discussed so far are listed below, each with its PMF/PDF. Each of these distributions is important because it has a natural, useful story, so understanding these stories (and recognizing equivalent stories) is crucial. It is also important to know how these distributions are related to each other. For example, $\text{Bern}(p)$ is the same as $\text{Bin}(1, p)$, and $\text{Bin}(n, p)$ is approximately $\text{Pois}(\lambda)$ if n is large, p is small, and $\lambda = np$ is moderate.

Name	Param.	PMF or PDF
Bernoulli	p	$P(X = 1) = p, P(X = 0) = 1 - p$
Binomial	n, p	$\binom{n}{k} p^k (1 - p)^{n-k}$, for $k \in \{0, 1, \dots, n\}$
Geometric	p	$(1 - p)^k p$, for $k \in \{0, 1, 2, \dots\}$
Negative Binomial	r, p	$\binom{r+n-1}{r-1} p^r (1 - p)^n$, $n \in \{0, 1, 2, \dots\}$
Hypergeometric	w, b, n	$\frac{\binom{w}{k} \binom{b}{n-k}}{\binom{w+b}{n}}$, for $k \in \{0, 1, \dots, n\}$
Poisson	λ	$\frac{e^{-\lambda} \lambda^k}{k!}$, for $k \in \{0, 1, 2, \dots\}$
Exponential	λ	$\lambda e^{-\lambda x}$ for $x > 0$
Uniform	$a < b$	$\frac{1}{b-a}$, for $x \in (a, b)$
Normal	μ, σ^2	$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

Some Useful Formulas

De Morgan's Laws

$$(A_1 \cup A_2 \cup \dots \cup A_n)^c = A_1^c \cap A_2^c \cap \dots \cap A_n^c$$

$$(A_1 \cap A_2 \cap \dots \cap A_n)^c = A_1^c \cup A_2^c \cup \dots \cup A_n^c$$

Complements

$$P(A^c) = 1 - P(A)$$

Unions

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

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = \sum_{i=1}^n P(A_i), \text{ if the } A_i \text{ are disjoint}$$

$$P(A_1 \cup A_2 \cup \dots \cup A_n) \leq \sum_{i=1}^n P(A_i)$$

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = \sum_{k=1}^n \left((-1)^{k+1} \sum_{i_1 < i_2 < \dots < i_k} P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) \right)$$

(Inclusion-Exclusion)

Intersections

$$P(A \cap B) = P(A)P(B|A) = P(B)P(A|B)$$

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1, A_2) \cdots P(A_n|A_1, \dots, A_{n-1})$$

Law of Total Probability

If $E_1, E_2, \dots, E_n$ are a partition of the sample space S (i.e., they are disjoint and their union is all of S) and $P(E_j) \neq 0$ for all j , then

$$P(B) = \sum_{j=1}^n P(B|E_j)P(E_j)$$

PROBABILITY AND STATISTICS (MATH-235)

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Mock exam - questions

Date: 18th of December, 2024

Time: 13:15–15:15

Name: _____

SCIPER: _____

INSTRUCTIONS TO CANDIDATES

- This is a **mock exam**. To obtain the maximum number of points you should be clear about your reasoning and present your arguments explicitly. You have **2 hours** to complete the **mock exam**, however for the **final exam** you will have **3 hours**. Accordingly, the number of questions will increase in the **final exam**.
- All that can be used for this exam is a pen and the cheat sheet that **we provide for you**. No external notes, books, summaries, formula collections or calculators are allowed. All questions should be answered.
- The finest enumerated item in each question will be marked on a scale of 0 – 2 points, indicating an incorrect, partially correct and completely correct answer respectively (half-points are not given). **The exam has 6 questions with a total of 54 points.**
- ~~Write the answer to every question in the other booklet (final exam – answers)~~ (*We do not have an answer booklet for the mock exam, one will be given at the final exam*). Scrap paper will be provided for rough work, but only answers written in the booklet will be marked.
- At the end of the exam, you will have to return everything : the booklet with the questions, the booklet with your answers, and the scrap paper.
- When the exercise directs you to “**Simplify!**” provide a simple fraction, a decimal, or the simplest form possible using the function or parameter in terms of which your answer must be given. If you are explicitly instructed with “**Do not simplify!**”, it is acceptable to leave your answer as an integral, sum, or another more complex form, depending on the context.

Mark question 1 (TOT: 8 points):

Mark question 2 (TOT: 10 points):

Mark question 3 (TOT: 6 points):

Mark question 4 (TOT: 12 points):

Mark question 5 (TOT: 6 points):

Mark question 6 (TOT: 12 points):

VARIABLE DEFINITIONS

All the random variables have finite mean and variance.

Question 1.

Two coins are in a hat. The coins look alike, but one coin is fair (with probability $1/2$ of Heads), while the other coin is biased, with probability $1/4$ of Heads. One of the coins is randomly pulled from the hat, without knowing which of the two it is. Call the chosen coin “Coin C”.

- (a) Coin C is tossed twice, showing Heads both times. Given this information, what is the probability that Coin C is the fair coin? (Simplify.)
- (b) Now suppose that coin C is tossed twice, showing Heads for the first time and Tails for the second. Given this information, what is the probability that Coin C is the fair coin? (Simplify.)
- (c) Are the events “first toss of Coin C is Heads” and “second toss of Coin C is Heads” independent? Explain briefly.
- (d) Find the probability that in 10 flips of Coin C, there will be exactly 3 Heads. (The coin is equally likely to be either of the 2 coins; do not assume it already landed Heads twice as in (a) or that it landed Heads then Tails as in (b). Do not simplify.)

Question 2.

- (a) Let $X_1, X_2, \dots$ be independent $\mathcal{N}(0, 4)$ random variables, and let J be the smallest value of j such that $X_j > 4$ (i.e., the index of the first X_j exceeding 4).
- (i) What is the distribution of $J - 1$? Derive the parameter(s) of it in terms of a probability, using the X_j -s!
 - (ii) In terms of the standard Normal CDF Φ , find $E(J)$ (simplify).
- (b) Let f and g be PDFs with $f(x) > 0$ and $g(x) > 0$ for all x . Let X be a random variable with PDF f . Find the expected value of the ratio $\frac{g(X)}{f(X)}$ (simplify).
- (c) Define $F(x) = e^{-e^{-x}}$. This is a CDF (called the Gumbel distribution) and is a continuous, strictly increasing function. Let X have CDF F , and define $W = F(X)$.
- (i) Find the CDF of W .
 - (ii) What are the mean and variance of W (simplify)?

Question 3.

Let $U \sim \text{Unif}(0, 1)$, and $X = \ln\left(\frac{U}{1-U}\right)$.

(a) Write down (but do not compute) an integral giving $E(X^2)$.

(b) Find the CDF of X (simplify).

(c) Find $E(X)$ *without using calculus* (simplify).

Hint: $1 - U$ has the same distribution as U .

Question 4.

Let X and Y be positive random variables, *not necessarily independent*. Assume that the various expected values below exist. Write the most appropriate of $\leq$, $\geq$, $=$, or $?$ in your answer booklet for each part (where “?” means that no relation holds in general). Write down the name of the rules, theorems or inequalities you used.

(a) $(E(XY))^2$ _____ $E(X^2)E(Y^2)$

(b) $P(|X + Y| > 2)$ _____ $\frac{1}{10}E((X + Y)^4)$

(c) $E(\ln(X + 3))$ _____ $\ln(E(X + 3))$

(d) $E(X^2e^X)$ _____ $E(X^2)E(e^X)$

(e) $P(X + Y = 2)$ _____ $P(X = 1)P(Y = 1)$

(f) $P(X \geq 1 \text{ or } Y \geq 1)$ _____ $P(\{X \geq 1\} \cup \{Y \geq 1\})$

Question 5.

- (a) A woman is pregnant with twin boys. Twins may be either identical or fraternal (non-identical). In general, $1/3$ of twins born are identical. Obviously, identical twins must be of the same sex; fraternal twins may or may not be. Assume that identical twins are equally likely to be both boys or both girls, while for fraternal twins all possibilities are equally likely. Given the above information, what is the probability that the woman's twins are identical?
- (b) A certain genetic characteristic is of interest. For a random person, this has a numerical value given by a $\mathcal{N}(0, \sigma^2)$ random variable. Let X_1 and X_2 be the values of the genetic characteristic for the twin boys from (a). If they are identical, then $X_1 = X_2$; if they are fraternal, then X_1 and X_2 have correlation ρ .
- (i) Find the value of $E(X_2^2)$ in terms of σ^2 .
- (ii) Find $\text{Cov}(X_1, X_2)$ in terms of ρ, σ^2 .

Question 6.

Cassie enters a casino with $X_0 = 1$ dollar and repeatedly plays the following game: with probability $1/3$, the amount of money she has increases by a factor of 3; with probability $2/3$, the amount of money she has decreases by a factor of 3. Let X_n be the amount of money she has after playing this game n times. For example, X_1 is 3 with probability $1/3$ and is 3^{-1} with probability $2/3$.

- (a) Compute $E(X_1)$, $E(X_2|X_1 = \frac{1}{3})$ and $E(X_2|X_1 = 3)$.
- (b) Compute $E(X_2)$ and, in general, $E(X_n)$.
- (c) What happens to $E(X_n)$ as $n \rightarrow \infty$?
- (d) Let Y_n be the number of times out of the first n games that Cassie triples her money. What happens to Y_n/n as $n \rightarrow \infty$? Which Theorem justifies your answer?
- (e) Find an expression for X_n in terms of Y_n .
- (f) What happens to X_n as $n \rightarrow \infty$?

Unofficial French Translation

Question 1.

Deux pièces sont placées dans un chapeau. Les pièces se ressemblent, mais l'une d'entre elles n'est pas biaisée (avec une probabilité de $1/2$ de faire face), tandis que l'autre est biaisée, avec une probabilité de $1/4$ de faire face. On tire au hasard l'une des pièces du chapeau, sans savoir de laquelle des deux il s'agit. On appelle la pièce choisie "Pièce C".

- (a) La pièce C est jetée deux fois et donne face à chaque fois. Compte tenu de ces informations, quelle est la probabilité que la pièce C soit la bonne ? (Simplifier.)
- (b) Supposons maintenant que la pièce C soit lancée deux fois et qu'elle donne face la première fois et pile la seconde. Compte tenu de cette information, quelle est la probabilité que la pièce C soit la bonne ? (Simplifier.)
- (c) Les événements "premier lancer de la pièce C est face" et "deuxième lancer de la pièce C est face" sont-ils indépendants ? Expliquez brièvement.
- (d) Trouvez la probabilité que sur 10 lancers de la pièce C, il y ait exactement 3 faces. (La pièce a la même probabilité d'être l'une ou l'autre des deux pièces ; ne supposez pas qu'elle est déjà tombée deux fois sur face comme en (a) ou qu'elle est tombée sur face puis sur pile comme en (b). Ne simplifiez pas).

Question 2.

- (a) Soit $X_1, X_2, \dots$ des variables aléatoires indépendantes $\mathcal{N}(0, 4)$, et soit J la plus petite valeur de j telle que $X_j > 4$ (c'est-à-dire que l'indice du premier X_j qui dépasse 4).
 - (i) Quelle est la distribution de $J - 1$? En déduire le(s) paramètre(s) en termes de probabilité, en utilisant les X_j -s !
 - (ii) En termes de la fonction de distribution cumulative (CDF) normale standard Φ , trouver $E(J)$ (simplifier).
- (b) Soit f et g des fonctions de densité de probabilités (PDFs) avec $f(x) > 0$ et $g(x) > 0$ pour tout x . Soit X une variable aléatoire de fonction de densité de probabilité (PDF) f . Trouver l'espérance du rapport $\frac{g(X)}{f(X)}$ (simplifier).
- (c) Définissez $F(x) = e^{-e^{-x}}$. Il s'agit d'une fonction de distribution cumulative (CDF) (appelée distribution de Gumbel) et d'une fonction continue, strictement croissante. Soit X ayant une fonction de distribution cumulative (CDF) F , et définissons $W = F(X)$.
 - (i) Trouvez la fonction de distribution cumulative (CDF) de W .
 - (ii) Quelles sont la moyenne et la variance de W (simplifier) ?

Question 3.

Soit $U \sim \text{Unif}(0, 1)$, et $X = \ln\left(\frac{U}{1-U}\right)$.

- (a) Ecrire (sans la calculer) une intégrale donnant $E(X^2)$.
- (b) Trouvez la fonction de distribution cumulative (CDF) de X (simplifiez).
- (c) Trouver $E(X)$ *sans utiliser le calcul* (simplifier). *Heure:* $1-U$ a la même distribution que U .

Question 4.

Soit X et Y des variables aléatoires positives, *pas nécessairement indépendantes*. Supposons que les différentes valeurs attendues ci-dessous existent. Pour chaque partie, écrivez dans votre livret de réponses la valeur la plus appropriée parmi $\leq$, $\geq$, $=$, ou $?$ (où “?” signifie qu’aucune relation n’existe en général). Notez le nom des règles, théorèmes ou inégalités que vous avez utilisés.

- (a) $(E(XY))^2$ ____ $E(X^2)E(Y^2)$
- (b) $P(|X + Y| > 2)$ ____ $\frac{1}{10}E((X + Y)^4)$
- (c) $E(\ln(X + 3))$ ____ $\ln(E(X + 3))$
- (d) $E(X^2e^X)$ ____ $E(X^2)E(e^X)$
- (e) $P(X + Y = 2)$ ____ $P(X = 1)P(Y = 1)$
- (f) $P(X \geq 1 \text{ ou } Y \geq 1)$ ____ $P(\{X \geq 1\} \cup \{Y \geq 1\})$

Question 5.

- (a) Une femme est enceinte de jumeaux. Les jumeaux peuvent être identiques ou frateruels (non identiques). En général, $1/3$ des jumeaux nés sont identiques. Il est évident que les vrais jumeaux doivent être du même sexe ; les faux jumeaux peuvent l’être ou non. Supposons que les vrais jumeaux ont la même probabilité d’être tous les deux garçons ou toutes les deux filles, alors que pour les faux jumeaux, toutes les possibilités sont également probables. Compte tenu des informations ci-dessus, quelle est la probabilité que les jumeaux de la femme soient identiques ?
- (b) On s’intéresse à une certaine caractéristique génétique. Pour une personne aléatoire, cette caractéristique a une valeur numérique donnée par une variable aléatoire $\mathcal{N}(0, \sigma^2)$. Soit X_1 et X_2 les valeurs de la caractéristique génétique pour les jumeaux de (a). S’ils sont identiques, alors $X_1 = X_2$; s’ils sont frateruels, alors X_1 et X_2 ont une corrélation ρ .
 - (i) Trouver la valeur de $E(X_2^2)$ en termes de σ^2 .
 - (ii) Trouver $\text{Cov}(X_1, X_2)$ en termes de ρ, σ^2 .

Question 6.

Cassie entre dans un casino avec $X_0 = 1$ dollar et joue plusieurs fois au jeu suivant : avec une probabilité de $1/3$, la somme d’argent qu’elle possède augmente d’un facteur de 3 ; avec une probabilité de $2/3$, la somme d’argent qu’elle possède diminue d’un facteur de 3. Soit

X_n la somme d'argent qu'elle possède après avoir joué n fois à ce jeu. Par exemple, X_1 vaut 3 avec une probabilité de $1/3$ et vaut 3^{-1} avec une probabilité de $2/3$.

- (a) Calculer $E(X_1)$, $E(X_2|X_1 = \frac{1}{3})$ et $E(X_2|X_1 = 3)$.
- (b) Calculer $E(X_2)$ et, en général, $E(X_n)$.
- (c) Qu'arrive-t-il à $E(X_n)$ lorsque $n \rightarrow \infty$?
- (d) Soit Y_n le nombre de fois sur les n premières parties où Cassie triple son argent. Qu'arrive-t-il à Y_n/n lorsque $n \rightarrow \infty$? Quel théorème justifie votre réponse ?
- (e) Trouvez une expression pour X_n en termes de Y_n .
- (f) Que se passe-t-il pour X_n lorsque $n \rightarrow \infty$?