¡article¿
¡presentation¿ Outline
Discrete random variables
Introduction
This material comes primarily from [][Chapter 4]rice07.
We will cover the ideas of expected value, variance, as well has higher-order moments.
This includes topics such as conditional expectation, which is one of the fundamental ideas behind many branches of sta
For instance, most regression / prediction algorithms are built with the idea of minimizing some conditional expectation
[]allowframebreaks]Expectation: Discrete random variables Definition: Expectation of discrete random variables Let X b

provided that $\sum_{x\in\mathcal{X}} |x|\, p(x) < \infty$; otherwise, the expectation is not defined.
This is not the most mathematically precise definition of expectation, but a more complete treatment of the topic is out
The concept of the expected value parallels the notion of a *weighted average*.
That is, we weight each possibility $x \in \mathcal{X}$ by their corresponding probability: $\sum_x x\, p(x)$.

Combining the factors of x in the integrand, we obtain

Now we will apply the “integration by density function” trick: we will re-write the integrand so that it corresponds to t
$= \int_0^\infty \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha^*-1} e^{-\lambda x}\, dx$
$= \left(\frac{\lambda^\alpha}{\Gamma(\alpha)}\right)\left(\frac{\Gamma(\alpha^*)}{\lambda^{\alpha^*}}\right) \int_0^\infty \frac{\lambda^{\alpha^*}}{\Gamma(\alpha^*)} x^{\alpha*-1} e^{-\lambda x}\, dx$
$= \left(\frac{\lambda^\alpha}{\Gamma(\alpha)}\right)\left(\frac{\Gamma(\alpha^*)}{\lambda^{\alpha^*}}\right)$ Where the last step is a result of the fact that the integrand (and support of the integral) matches th
¡presentation¿ *Solution.*
Expectation of functions of random variables
allowframebreaks]Functions of random variables
We are often interested in functions of random variables: $Y = g(X)$.
Ideas that we have already covered enable us to calculate $E(Y)$.
For instance, you could use the change-of-variables theorem to get the density of Y, then use the definition to calculate
Fortunately, we don't have to do this. We can instead calculate $E[Y]$ by integrating (or summing) with respect to X:

We will justify this for the discrete case.
Theorem 4.1: Expectation of transformed random variables Suppose that X is a random variable and that $Y = g(X)$ fo
If X is discrete with pmf $p(x)$:

provided that $\sum_x |g(x)| p(x) < \infty$.
If X is continuous with pdf $f(x)$:

provided that $\int |g(x)| f(x)\, dx < \infty$.
allowframebreaks]Functions of random variables: proof
¡presentation¿ *Proof:*
¡article¿ *Proof:* By definition of expectation,

Now let A_i denote the set of x's that are mapped to y_i by g. That is, A_i is the pre-image of y_i, meaning that $x \in A_i$ if

and we can express the expectation as $\mathrm{E(Y)} = \sum_i y_i p_Y(y_i)$
$= \sum_i y_i \sum_{x \in A_i} p(x)$
$= \sum_i \sum_{x \in A_i} y_i\, p(x)$
$= \sum_i \sum_{x \in A_i} g(x)\, p(x)$
$= \sum_x g(x)\, p(x)$ Here, the second to last step is because for all $x \in A_i$, $g(x) = y_i$ by definition. The final step is a result
The proof for the continuous case is similar, but does require a measure-theoretic approach to integration.
One important thing to note is that $g\big(E(X)\big)$ is not usually equal to $E\big(g(x)\big)$.
For example, let Z be a standard normal. We know that $E[Z] = 0$, because it's symmetric. However, $P\big(|Z| > 0\big) = 1$, t
This idea can be extended to show that if for all non-negative random variables X that have finite expectation, if $g(x)$
allowframebreaks]Expected value of indicator functions
Another important example of expectations is indicator random variables.
For example, suppose that X is a random variable. Then $Y = 1[X \in A]$ for some $A \subset \mathcal{X}$ is a random variable.
Indicator Random Variable Let X follow a standard normal distribution, and $A = [-1, 1]$. Then $Y = 1[X \in A]$ is define
Expectations of indicator variables are probabilities. Let $Y = 1[X \in A]$. $\mathrm{E(Y)} = \mathrm{E}\big(1[\mathrm{X} \in A]\big)$
$= \int_{x \in \mathcal{X}} 1[X \in A]\, f(x)\, dx$
$= \int_{x \in A} f(x)\, dx = P(X \in A)$.
This fact is useful for deriving some important inequalities.
First, we will show that the expectations of interest actually exist.
Let X be a continuous random variable with expectation $E(X)$. From our definition, this implies that $\int |x|\, f(x)\, dx < \infty$
Now suppose that for some random variable $Y = g(X)$ such that $|Y| \le |X|$. Then we can deduce that $\int |y|\, f(x)\, dx < \infty$
Now suppose that φ is a non-decreasing, non-negative function, and that for some $a \in R$, $\varphi(a) > 0$. Then, for all $x \ge a$,
Define $Y = 1[X \ge a]$. Note that for all possible outcomes $\omega \in \Omega$,

Taking expectations of everything (which we argued preserves inequalities),

This inequality is known as Markov's (general) inequality, and is very useful for bounding the probability of particular

Acknowledgments
Compiled on using Rversion 4.5.1.

We acknowledge https://jeswheel.github.io/4450$_f$25/$acknowledge.htmlstudentsandinstructorsforpreviousversionsoft$