

# Important Results - MATH 356

---

Ralph Sarkis 260729917

March 14, 2019

## 1 Probability Basics

### Proposition 1.1.

- $\mathbb{P}(\emptyset) = 0$  because  $\mathbb{P}(\emptyset) = \mathbb{P}(\bigcup_{i \in \mathbb{N}} \emptyset) = \sum_{i \in \mathbb{N}} \mathbb{P}(\emptyset)$  using countable additivity.
- If  $A_i \in \mathcal{S}$  for  $i = 1, \dots, n$  with  $A_i \cap A_j = \emptyset$  for all  $i \neq j$ , we have  $\mathbb{P}(\bigcap_{i=1}^n A_i) = \sum_{i=1}^n \mathbb{P}(A_i)$  (finite additivity)
- $\forall A \in \mathcal{S}, \mathbb{P}(A) \in [0, 1]$ , the proof of that also shows that  $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$
- If  $A, B \in \mathcal{S}$  and  $A \subseteq B$ , we have  $\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A)$ , in particular, we have  $\mathbb{P}(A) \leq \mathbb{P}(B)$  which is called monotonicity
- Let  $A, B \in \mathcal{S}$ , we have  $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$  (finite sub-additivity)
- Let  $A_i \in \mathcal{S}$  for  $i = 1, \dots, n$ , the Inclusion-Exclusion formula says the following:

$$\mathbb{P}\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n \mathbb{P}(A_i) - \sum_{\substack{i_1, i_2=1 \\ i_1 < i_2}}^n \mathbb{P}(A_{i_1} \cap A_{i_2}) + \dots + (-1)^{n+1} \mathbb{P}\left(\bigcap_{i=1}^n A_i\right)$$

**Theorem 1.2** (Bonferonni's inequality). Let  $A_i \in \mathcal{S}$  for  $i = 1, \dots, n$ , we have the following inequality:

$$\sum_{i=1}^n \mathbb{P}(A_i) - \sum_{\substack{i_1, i_2=1 \\ i_1 < i_2}}^n \mathbb{P}(A_{i_1} \cap A_{i_2}) \leq \mathbb{P}\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n \mathbb{P}(A_i)$$

**Proposition 1.3.** Let  $\mathcal{S}$  be a  $\sigma$ -field and  $\mathbb{P}$  a probability distribution.

- *Continuity from below* : If  $A_n \in \mathcal{S}$  for  $n \in \mathbb{N}$  and  $A_n \uparrow A$ , then, the sequence  $\mathbb{P}(A_n)$  is monotonically increasing and  $\lim_{n \rightarrow \infty} \mathbb{P}(A_n) = \mathbb{P}(A)$ .
- *Continuity from above* : If  $A_n \in \mathcal{S}$  for  $n \in \mathbb{N}$  and  $A_n \downarrow A$ , then, the sequence  $\mathbb{P}(A_n)$  is monotonically decreasing and  $\lim_{n \rightarrow \infty} \mathbb{P}(A_n) = \mathbb{P}(A)$ .
- *Countable sub-additivity* : If  $A_n \in \mathcal{S}$  for  $n \in \mathbb{N}$ , we have  $\mathbb{P}(\bigcup_{n \in \mathbb{N}} A_n) \leq \sum_{n=1}^{\infty} \mathbb{P}(A_n)$ .

## 2 Conditional Probability

**Definition 2.1** (Conditional probability). Let  $(\Omega, \mathcal{S}, \mathbb{P})$  be a probability space and  $H \in \mathcal{S}$  such that  $\mathbb{P}(H) > 0$ . Then, for every  $A \in \mathcal{S}$ , we define the conditional probability of  $A$  given  $H$  by

$$\mathbb{P}(A|H) = \frac{\mathbb{P}(A \cap H)}{\mathbb{P}(H)}$$

If  $\mathbb{P}(H) = 0$ , then  $\mathbb{P}(A|H)$  is undefined.

**Theorem 2.2** (The multiplication rule). If  $A_i \in \mathcal{S}$  for  $i \in \{1, \dots, n\}$  and  $\mathbb{P}(\bigcap_{i=1}^{n-1} A_i) > 0$ , then

$$\mathbb{P}\left(\bigcap_{i=1}^n A_i\right) = \mathbb{P}(A_1)\mathbb{P}(A_2|A_1)\mathbb{P}(A_3|A_1 \cap A_2) \cdots \mathbb{P}\left(A_n \mid \bigcap_{j=1}^{n-1} A_j\right)$$

**Theorem 2.3** (Total probability and Bayes' rule). Let  $\{H_i \mid i \geq 1\}$  be a partition of  $\Omega$ , namely, they are disjoint sets whose union is equal to  $\Omega$ , such that  $\forall i \geq 1, \mathbb{P}(H_i) > 0$ . Then, the total probability rule states that  $\forall A \in \mathcal{S}, \mathbb{P}(A) = \sum_{i=1}^{\infty} \mathbb{P}(A|H_i)\mathbb{P}(H_i)$ . In particular, Bayes' rule can be derived :

$$\mathbb{P}(H_i|A) = \frac{\mathbb{P}(A|H_i)\mathbb{P}(H_i)}{\sum_{j=1}^{\infty} \mathbb{P}(A|H_j)\mathbb{P}(H_j)}$$

**Definition 2.4** (Lack of memory). Let  $(\Omega, \mathcal{S}, \mathbb{P})$  be a discrete probability space with  $\Omega = \mathbb{N}$  and  $\mathcal{S} = \mathcal{P}(\Omega)$ . We say that  $\mathbb{P}$  is memoryless if for any  $m, k \in \mathbb{N}$ , we have  $\mathbb{P}(\{k+m\} \mid \{m, m+1, \dots\}) = f(k)$ . One can show that  $\mathbb{P}$  must be a geometric distribution. Moreover, in the continuous case,  $\mathbb{P}$  must be an exponential distribution.

## 3 Random Variables

**Definition 3.1** (PMF). Let  $\mathbb{P}$  be a discrete distribution on some discrete probability space. The function  $f : \Omega \rightarrow [0, 1]$  defined with  $f(\omega) = \mathbb{P}(\{\omega\})$  is called the probability mass function. It has the property that  $\sum_{\omega \in \Omega} f(\omega) = 1$ . Moreover, if a function has this property, then it is the PMF of some discrete distribution.

We now give some examples of discrete distributions which we will use often because they describe what we usually study with probability. Note that we most often work with the probability space  $(\Omega, \mathcal{S}, \mathbb{P})$  and denote the PMF with  $f$ .

**Definition 3.2** (Discrete Uniform). If  $|\Omega| = n < \infty$  and the PMF of  $\mathbb{P}$  is  $f(\omega) = \frac{1}{n}$ ,  $\mathbb{P}$  is called the discrete uniform distribution. Let's verify that  $f$  is indeed a PMF.

$$\sum_{\omega \in \Omega} f(\omega) = \sum_{\omega \in \Omega} \frac{1}{n} = n \cdot \frac{1}{n} = 1$$

This distribution is used to describe experiences where each sample point happens as often as the others. For example, a fair coin/dice, three dices, etc...

**Definition 3.3** (Dirac). If the PMF of  $\mathbb{P}$  is such that there exists a sample point  $\omega_0 \in \Omega$  such that  $f(\omega_0) = 1$ , we call  $\mathbb{P}$  a Dirac distribution. We can infer that all the other sample points have probability zero. Let's verify that  $f$  is indeed a PMF.

$$\sum_{\omega \in \Omega} f(\omega) = \sum_{\omega \in \Omega \setminus \{\omega_0\}} 0 + f(\omega_0) = 1$$

This distribution is used when only one possible outcome happens all the time.

**Definition 3.4** (Bernoulli). If  $\Omega = \{0, 1\}$  and there is a number  $p \in (0, 1)$  such that  $f(0) = 1 - p$  and  $f(1) = p$ , we call  $\mathbb{P}$  the Bernoulli distribution and denote it  $B(p)$ . Let's verify that  $f$  is indeed a PMF.

$$\sum_{\omega \in \Omega} f(\omega) = f(1) + f(0) = p + 1 - p = 1$$

This distribution is used when the experience has two outcomes, where one has probability  $p$  of happening. We often call this outcome the success and the other outcome the failure, so the experience has probability  $p$  of succeeding.

**Definition 3.5** (Binomial). If  $\Omega = \{0, \dots, n\}$  for  $n > 0$  and there is a number  $p \in (0, 1)$  such that  $f(k) = \binom{n}{k} p^k (1 - p)^{n-k}$  for every  $k \in \Omega$ , we call  $\mathbb{P}$  the Binomial distribution and denote it  $B(n, p)$ . Let's verify that  $f$  is indeed a PMF<sup>1</sup>.

$$\sum_{\omega \in \Omega} f(\omega) = \sum_{k=0}^n \binom{n}{k} p^k (1 - p)^{n-k} = (p + (1 - p))^n = 1$$

This distribution is used to describe the probability of having  $k$  successes in  $n$  independent Bernoulli trials. In particular  $B(1, p) = B(p)$ . A concrete example of that is the probability of getting 4 heads when flipping 5 weighted coins that fall on head 75% of the time.

---

<sup>1</sup>We will use the binomial formula

**Definition 3.6 (Poisson).** If  $\Omega = \mathbb{N}$  and there is a number  $\lambda > 0$  such that  $f(k) = e^{-\lambda} \frac{\lambda^k}{k!}$  for every  $k \in \mathbb{N}$ , we call  $\mathbb{P}$  the Poisson distribution and denote it  $P(\lambda)$ . Let's verify that  $f$  is indeed a PMF<sup>2</sup>.

$$\sum_{\omega \in \Omega} f(\omega) = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} e^{\lambda} = 1$$

This distribution models the number of rare events occurring in a certain period when we know that the rare events have a constant rate and occur independently from the last time since the last event. It can also be viewed as the limit of  $B(n, \frac{\lambda}{n})$  as  $n \rightarrow \infty$ .

**Definition 3.7 (Geometric).** If  $\Omega = \mathbb{N}$  and there is a number  $p \in (0, 1)$  such that  $f(k) = (1 - p)^k p$  for every  $k \in \mathbb{N}$ , we call  $\mathbb{P}$  the Geometric distribution and denote it  $G(p)$ . Let's verify that  $f$  is indeed a PMF<sup>3</sup>.

$$\sum_{\omega \in \Omega} f(\omega) = \sum_{k=0}^{\infty} (1 - p)^k p = p \sum_{k=0}^{\infty} (1 - p)^k = p \frac{1}{1 - (1 - p)} = 1$$

This distribution models the number of failed Bernoulli trials before the first success.

Although the PMF gives all the information needed for a discrete distribution, it is not enough for a continuous distribution. Hence, we will explore a new notion.

**Definition 3.8 (CDF).** Consider the probability space  $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{Q})$ . The cumulative distribution function  $F$  of  $\mathbb{Q}$  is defined as  $F : \mathbb{R} \rightarrow [0, 1]$  and  $F(x) = \mathbb{Q}((-\infty, x])$ . It has three important properties :

- $F$  is non-decreasing on  $\mathbb{R}$ .
- $\lim_{x \rightarrow -\infty} F(x) = 0$  and  $\lim_{x \rightarrow \infty} F(x) = 1$
- $F$  is right continuous.

Moreover, if some function  $F$  satisfies these three properties, then it is the cumulative distribution function for some distribution on  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ .

In order to study their CDFs, we will look at discrete distributions on the space  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ .

**Example 3.9.**

- Consider the distribution  $B(n, p)$  on  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ , here is its CDF.

$$F(x) = \mathbb{P}((-\infty, x]) = \sum_{k \leq x} f(k) = \sum_{k \leq x} \binom{n}{k} p^k (1 - p)^{n-k}$$

Here are some values :  $x < 0 \implies F(x) = 0$ ,  $0 \leq x \leq 1 \implies F(x) = (1 - p)^n$ ,  
 $1 \leq x \leq 2 \implies F(x) = (1 - p)^n + np(1 - p)^{n-1}$ , etc...

---

<sup>2</sup>We use the Taylor expansion of the function  $e^x$

<sup>3</sup>We use the formula of the geometric series

- Consider the distribution  $P(\lambda)$  on  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ , here is its CDF.

$$F(x) = \mathbb{P}((-\infty, x]) = \sum_{k \leq x} f(k) = \sum_{k \leq x} e^{-\lambda} \frac{\lambda^k}{k!}$$

Here are some values :  $x < 0 \implies F(x) = 0$ ,  $0 \leq x \leq 1 \implies F(x) = e^{-\lambda}$ ,  
 $1 \leq x \leq 2 \implies F(x) = e^{-\lambda} + \lambda e^{-\lambda}$ , etc...

- Consider the distribution  $G(p)$  on  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ , here is its CDF.

$$F(x) = \mathbb{P}((-\infty, x]) = \sum_{k \leq x} f(k) = \sum_{k \leq x} (1-p)^k p$$

Here are some values :  $x < 0 \implies F(x) = 0$ ,  $0 \leq x \leq 1 \implies F(x) = p$ ,  
 $1 \leq x \leq 2 \implies F(x) = p(1-p)$ , etc...

The CDF does not give more information than the PMF for discrete distributions, so we will look at continuous distribution. The simplest example being the uniform distribution.

**Definition 3.10** (Continuous uniform). Let  $Q$  be a distribution on  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$  and  $a < b$  be two real numbers such that the CDF of  $Q$  is the following :

$$F(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & x \in [a, b] \\ 1 & x > b \end{cases}$$

We call  $Q$  the uniform distribution on  $[a, b]$  and we denote it  $U(a, b)$ . Let's verify that  $F$  is indeed a CDF. It is clearly non-decreasing since each part is non-decreasing and  $\forall x \in [a, b], \frac{x-a}{b-a} \in [0, 1]$ . Moreover, we have  $\lim_{x \rightarrow -\infty} F(x) = 0$  and  $\lim_{x \rightarrow \infty} F(x) = 1$  since  $x \leq a \implies F(x) = 0$  and  $x \geq b \implies F(x) = 1$ . Lastly, we want to show that it is right continuous, inside the intervals, this is clear but we need to check at  $x = a$  and  $x = b$ <sup>4</sup>.

$$\lim_{x \rightarrow a^+} F(x) = \lim_{n \rightarrow \infty} F(a + \frac{1}{n}) = \lim_{n \rightarrow \infty} \frac{1}{n} \frac{1}{b-a} = 0 = F(a)$$

$$\lim_{x \rightarrow b^+} F(x) = \lim_{n \rightarrow \infty} F(b + \frac{1}{n}) = 1 = F(b)$$

We will now look at other continuous distributions and their definitions.

**Definition 3.11** (Continuous distribution). Consider the space  $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{Q})$  and the CDF  $F$ . We say that  $Q$  is a continuous distribution if  $F$  is a absolutely continuous, namely,  $\exists f : \mathbb{R} \rightarrow [0, \infty)$  such that  $\forall x \in \mathbb{R}, F(x) = \int_{-\infty}^x f(t) dt$ . We call  $f$  the probability density function (PDF). Here are some properties of  $Q$ .

<sup>4</sup>This is not really a proof since we are using only one sequence to show the limit but it can be refined

- $\int_{-\infty}^{\infty} f(t)dt = 1$
- $\forall a \leq b, \mathbf{Q}((a, b]) = F(b) - F(a) = \int_a^b f(t)dt$
- $F$  is continuous everywhere and differentiable almost everywhere. Also,  $F'(x) = f(x)$  at almost every  $x$  and at every  $x$  where  $f$  is continuous.

**Examples 3.12.**

- For  $U(a, b)$ , the PDF is  $f(x) = \frac{1}{b-a}$  for  $x \in [a, b]$  and  $f(x) = 0$  everywhere else.
- The Gaussian distribution (also called normal) is denoted  $N(0, 1)$  and has the following CDF and PDF.

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt \quad f(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$$

We will use these distribution to talk about random variables.

**Definition 3.13 (RV).** Let  $\Omega$  be a simple space and  $\mathcal{S}$  be a  $\sigma$ -field on  $\Omega$ . A function  $X : \Omega \rightarrow \mathcal{S}$  is called a random variable if  $\forall B \in \mathcal{B}(\mathbb{R}), X^{-1}(B) \in \mathcal{S}$ .

Note that the notion of probability is not part of this definition and if  $\Omega$  is discrete and  $\mathcal{S} = 2^\Omega$ , any function from  $\Omega$  to  $\mathbb{R}$  is a random variable. An equivalent definition is that  $\forall x \in \mathbb{R}, \{X \leq x\} = \{\omega \in \Omega \mid X(\omega) < x\} \in \mathcal{S}$ . Using this definition, it is easy to show that if  $X$  is a RV,  $aX + b, e^X$  and other simple transformations of  $X$  are RVs. More generally, if  $g$  is a Borel function (i.e:  $g : \mathbb{R} \rightarrow \mathbb{R}$  and  $\forall B \in \mathcal{B}(\mathbb{R}), g^{-1}(B) \in \mathcal{B}(\mathbb{R})$ ) then  $g \circ X$  is a RV as well.

Now, we introduce the notion of probability in RVs.

**Definition 3.14 (Distribution of RV).** Given a probability space  $(\Omega, \mathcal{S}, \mathbb{P})$  and a RV  $X$ . Define the following set function  $\mathbf{Q} : \mathcal{B}(\mathbb{R}) \rightarrow \mathbb{R}$  by  $B \mapsto \mathbb{P}(X^{-1}(B))$ . The CDF of  $\mathbf{Q}$  is also the CDF of  $X$  and it can be written as  $F(X) = \mathbb{P}(X \leq x)$ . If  $g$  is a Borel function, we have  $\mathbf{Q}_{g \circ X}(B) = \mathbf{Q}_X(g^{-1}(B))$ .

We now turn to expectation and variance of a RV.

**Definition 3.15 (Discrete expectation).** If  $X$  is a discrete RV on  $(\Omega, \mathcal{S}, \mathbb{P})$  and  $X \in \{x_n \mid n \in \mathbb{N}\}$  with PMF  $f_X$ . The expectation of  $X$ , denoted  $\mathbb{E}[X]$ , is defined if  $\sum_{n=0}^{\infty} |x_n| f_X(x_n) < \infty$  and we have  $\mathbb{E}[X] = \sum_{n=0}^{\infty} x_n f_X(x_n)$ .

**Proposition 3.16.** Here are some properties of the expectation. Let  $X$  and  $Y$  be RVs with  $\mathbb{E}[|X|] < \infty$  and  $\mathbb{E}[|Y|] < \infty$ , then the following holds.

- Let  $a, b \in \mathbb{R}$ , then  $\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$  (linearity).
- If  $X \geq Y$ , then  $\mathbb{E}[X] \geq \mathbb{E}[Y]$  (monotonicity).
- If  $\mathbb{E}[X^2]$  and  $\mathbb{E}[Y^2]$  also exist, then  $|\mathbb{E}[XY]| \leq \mathbb{E}[|XY|] \leq \sqrt{\mathbb{E}[X^2]\mathbb{E}[Y^2]}$ .

Let's calculate the expectations for various discrete RVs we know. We will assume that the expectation exists to avoid some computations.

**Examples 3.17.** We will work in the discrete probability space  $(\Omega, \mathcal{S}, \mathbb{P})$  and we will denote  $f$  to be the PMF of the RV.

- Let  $X$  be a Dirac RV with  $f(x_0) = 1$ , then  $\mathbb{E}[X] = x_0$ .
- Let  $X$  be a Bernoulli RV ( $X \in \{0, 1\}$ ) with parameter  $p \in (0, 1)$ , then  $\mathbb{E}[X] = 0 \cdot (1 - p) + 1 \cdot p = p$ .
- Let  $X$  be a Binomial RV ( $X \in \{0, n\}$ ) with parameter  $n > 0$  and  $p \in (0, 1)$ , we get the following.

$$\begin{aligned}
 \mathbb{E}[X] &= \sum_{k=0}^n k \cdot \binom{n}{k} p^k (1-p)^{n-k} && \text{(definition of expectation)} \\
 &= \sum_{k=1}^n k \cdot \binom{n}{k} p^k (1-p)^{n-k} && \text{(since first term is 0)} \\
 &= np \sum_{k=1}^n k \frac{(n-1)!}{(n-k)!k!} p^{k-1} (1-p)^{n-k} && \text{(factoring } np) \\
 &= np \sum_{k=1}^n \frac{(n-1)!}{(n-1-(k-1))!(k-1)!} p^{k-1} (1-p)^{n-1-(k-1)} && \text{(developping stuff)} \\
 &= np \sum_{j=0}^n \frac{(n-1)!}{(n-1-j)!j!} p^j (1-p)^{n-1-j} && \text{(putting } j = k-1) \\
 &= np && \text{(using the sum of PMF = 1)}
 \end{aligned}$$

- Let  $X$  be a Poisson RV ( $X \in \mathbb{N}$ ) with parameter  $\lambda > 0$ , we get the following.

$$\begin{aligned}
 \mathbb{E}[X] &= \sum_{k=0}^{\infty} k \cdot e^{-\lambda} \frac{\lambda^k}{k!} && \text{(definition of expectation)} \\
 &= \sum_{k=1}^{\infty} k \cdot e^{-\lambda} \frac{\lambda^k}{k!} && \text{(since the first term is 0)} \\
 &= \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} && \text{(factoring } \lambda e^{-\lambda} \text{ and simplifying)} \\
 &= \lambda e^{-\lambda} \sum_{j=0}^{\infty} \frac{\lambda^j}{(j)!} && \text{(set } j = k-1) \\
 &= \lambda e^{-\lambda} e^{\lambda} && \text{(using Taylor expansion of } e^x) \\
 &= \lambda && \text{(simplifying)}
 \end{aligned}$$

- Let  $X$  be a Geometric RV ( $X \in \mathbb{N}$ ) with parameter  $p \in (0, 1)$ , we get the following.

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{k=0}^n k \cdot (1-p)^k p && \text{(definition of expectation)} \\
&= \sum_{k=1}^n k \cdot (1-p)^k p && \text{(since first term is 0)} \\
&= \sum_{j=0}^n (j+1) \cdot (1-p)^{j+1} p && \text{(set } j = k - 1) \\
&= (1-p) \sum_{j=0}^n j \cdot (1-p)^j p + (1-p) \sum_{j=0}^n (1-p)^j p && \text{(develop the sum)} \\
&= (1-p)\mathbb{E}[X] + (1-p) && \text{(using the sum of PMF of } G(p))
\end{aligned}$$

Solving this for  $\mathbb{E}[X]$  gives us  $\mathbb{E}[X] = \frac{1-p}{p}$ .

**Definition 3.18** (Moments). If  $\mathbb{E}[X^k]$  is defined, it is called the  $k$ -th moment of  $X$ .

**Definition 3.19** (Variance). If the second moment of  $X$  is defined, the variance of  $X$  is also defined like so;

$$\mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

**Proposition 3.20.** Let  $X$  be a RV with  $\text{Var}(X) = \sigma < \infty$ , then  $\text{Var}(aX + b) = a^2\sigma$ .

**Definition 3.21** (MGF). If  $\exists s_0 > 0$  such that  $\forall s \in (-s_0, s_0)$ ,  $M(s) = \mathbb{E}[e^{sX}]$  is defined, then, the Moment Generating function is defined as  $\mathbb{E}[e^{sX}]$ , denoted  $M(s)$ . If  $M(s)$  exists, then  $\mathbb{E}[X^k]$  exists for all  $k \geq 0$  and  $M^{(k)}(0) = \mathbb{E}[X^k]$ .

**Proposition 3.22.** If  $X$  is a continuous RV with a defined MGF  $M(s)$ , assume that for any  $s \in \mathbb{R}$ ,  $M(2s) = M(s)^4$ ,  $\mathbb{E}[X] = 0$  and  $\text{Var}(X) = 1$ , then  $X$  is a  $N(0, 1)$  RV.

Let's calculate the MGF and variance for the same discrete RVs as above. Once again, we assume that they exist

**Examples 3.23.**

- Let  $X$  be Dirac RV with  $f(x_0) = 1$ . Then  $M(s) = \mathbb{E}[e^{sX}] = e^{sx_0}$ . We indeed get

$$M'(0) = x_0 e^{sx_0}|_{s=0} = x_0 = \mathbb{E}[X]$$

Moreover, we have

$$\mathbb{E}[X^2] = M''(0) = x_0^2 e^{sx_0}|_{s=0} = x_0^2$$

We can then calculate the variance<sup>5</sup>

$$\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = x_0^2 - x_0^2 = 0$$

---

<sup>5</sup>Note that the variance of a RV is 0 if and only if it is a Dirac RV

- Let  $X$  be a Bernoulli RV with parameter  $p \in (0, 1)$ , we find its MGF:

$$M(s) = \mathbb{E}[e^{sX}] = e^{s \cdot 0}(1 - p) + e^{s \cdot 1}(p) = (1 - p) + pe^s$$

Indeed, we find  $M'(0) = pe^s|_{s=0} = p = \mathbb{E}[X]$ . Moreover, we have

$$\mathbb{E}[X^2] = M''(0) = pe^s|_{s=0} = p$$

We can then calculate the variance

$$\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = p - p^2 = p(1 - p)$$

- Let  $X$  be a Binomial RV with parameter  $n > 0$  and  $p \in (0, 1)$ . We find its MGF:

$$\begin{aligned} M(s) = \mathbb{E}[e^{sX}] &= \sum_{k=0}^n e^{sk} \binom{n}{k} \cdot p^k (1 - p)^{n-k} \\ &= \sum_{k=0}^n \binom{n}{k} \cdot (pe^s)^k (1 - p)^{n-k} \\ &= (pe^s + 1 - p)^n \quad \text{(using the binomial formula)} \end{aligned}$$

Indeed, we get

$$M'(0) = n(pe^s + 1 - p)^{n-1}(pe^s)|_{s=0} = np = \mathbb{E}[X]$$

Moreover, we have

$$\begin{aligned} \mathbb{E}[X^2] &= M''(0) \\ &= npe^s(p(e^s - 1) + 1)^{n-2}(p(ne^s - 1) + 1)|_{s=0} \\ &= np(p(n - 1) + 1) \end{aligned}$$

So we can find the variance:

$$\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = np(p(n - 1) + 1) - (np)^2 = np(1 - p)$$

- Let  $X$  be a Poisson RV with parameter  $\lambda > 0$ . We can find its MGF.

$$\begin{aligned} M(s) = \mathbb{E}[e^{sX}] &= \sum_{k=0}^{\infty} e^{sk} e^{-\lambda} \cdot \frac{\lambda^k}{k!} \\ &= e^{-\lambda} \sum_{k=0}^{\infty} \frac{(e^s \lambda)^k}{k!} \\ &= e^{-\lambda} e^{e^s \lambda} \quad \text{(using the Taylor expansion of } e^x) \\ &= e^{\lambda(e^s - 1)} \end{aligned}$$

Indeed, we have  $M'(0) = \lambda e^s e^{\lambda(e^s-1)}|_{s=0} = \lambda$ . Moreover, we have

$$\begin{aligned}\mathbb{E}[X^2] &= M''(0) = (\lambda e^s + 1)(\lambda e^s e^{\lambda(e^s-1)})|_{s=0} \\ &= \lambda^2 + \lambda\end{aligned}$$

So we can find the variance

$$\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

- Let  $X$  be a Geometric RV with parameter  $p > 0$ . We find its MGF:

$$\begin{aligned}M(s) &= \mathbb{E}[e^{sX}] = \sum_{k=0}^{\infty} e^{sk} \cdot (1-p)^k p \\ &= p \sum_{k=0}^{\infty} (e^s(1-p))^k \\ &= p \cdot \frac{1}{1 - (e^s(1-p))}\end{aligned}$$

Indeed, we have

$$M'(0) = p \cdot \frac{(1-p)e^s}{(1 - e^s(1-p))^2}|_{s=0} = \frac{1-p}{p}$$

Moreover, we have

$$\mathbb{E}[X^2] = M''(0) = \frac{p(p-1)e^s((p-1)e^s - 1)}{((p-1)e^s + 1)^3}|_{s=0} = \frac{(p-1)(p-2)}{p^2}$$

So we can find the variance :

$$\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{(p-1)(p-2)}{p^2} - \left(\frac{1-p}{p}\right)^2 = \frac{1-p}{p^2}$$

## 4 Inequalities and Estimates

**Proposition 4.1.** *Let  $X \in \mathbb{N}$  be a discrete RV, then*

$$\mathbb{E}[X] = \sum_{i=1}^{\infty} \mathbb{P}(X \geq i)$$

**Proposition 4.2.** *Let  $X \in \mathbb{R}^+$  be a continuous and nonnegative RV, then*

$$\mathbb{E}[X] = \int_0^{\infty} \mathbb{P}(X > t) dt$$

**Definition 4.3.** For  $t > 0$ ,  $\mathbb{P}(X > t)$  is called the tail probability. This section is all about deriving inequalities and estimates on it.

First, we derive some sufficient (but not necessary) conditions for the existence of moments and generating functions (it works for any RV  $X$ ).

**Proposition 4.4.** *If there exists  $\delta > 0$  and  $C > 0$  such that for any  $t \in [0, \infty)$ ,  $P(|X| > t) \leq \frac{C}{t^{1+\delta}}$ , then*

$$\mathbb{E}[|X|] \leq \int_0^\infty \frac{C}{t^{1+\delta}} dt < \infty$$

**Proposition 4.5.** *Fix  $k \in \mathbb{N}$ , if there exists  $\delta > 0$  and  $C > 0$  such that for any  $t \in [0, \infty)$ ,  $P(|X|^k > t) \leq \frac{C}{t^{1+\frac{\delta}{k}}}$ , then*

$$\mathbb{E}[|X|^k] \leq \int_0^\infty \frac{C}{t^{1+\frac{\delta}{k}}} dt < \infty$$

**Proposition 4.6.** *If there exists  $\delta > 0$  and  $C > 0$  such that for any  $t \in [0, \infty)$ ,  $P(|X| > t) \leq Ce^{-\delta t}$ , then for any  $s \in (-\delta, \delta)$ , we have*

$$\mathbb{E}[e^{sX}] \leq \int_0^\infty tCe^{-\frac{\delta}{|s|}t} dt < \infty$$

**Theorem 4.7.** *Let  $X$  be a RV and  $h$  be a Borel function with  $h \geq 0$ . Assume that  $\mathbb{E}[h \circ X]$  exists, then for any  $t > 0$ , we have*

$$\mathbb{P}(h(X) > t) \leq \frac{\mathbb{E}[h(X)]}{t}$$

From this theorem, we can derive three important inequalities.

**Theorem 4.8** (Markov's inequality). *If  $X$  is a RV with  $\mathbb{E}[|X|^k] < \infty$ , then for any  $t > 0$ ,*

$$\mathbb{P}(|X| > t) = \mathbb{P}(|X|^k > t^k) \leq \frac{\mathbb{E}[|X|^k]}{t^k}$$

**Theorem 4.9** (Chebyshev's inequality). *If  $X$  is a RV with  $\mathbb{E}[|X|^2] < \infty$ , then for any  $t > 0$ ,*

$$\mathbb{P}(|X - \mathbb{E}[X]| > t) = \mathbb{P}(|X - \mathbb{E}[X]|^2 > t^2) \leq \frac{\text{Var}(X)}{t^2}$$

**Theorem 4.10** (Chernoff's bound). *If  $X$  is a RV with  $\mathbb{E}[e^{s|X|}] < \infty$ , then for any  $t > 0$ ,*

$$\mathbb{P}(|X| > t) = \mathbb{P}(e^{s|X|} > e^{st}) \leq e^{-st} \mathbb{E}[e^{s|X|}]$$

# 5 Multiple Random Variables

## 5.1 Distributions

**Definition 5.1** (Joint CDF). The joint CDF of a RV  $(X, Y)$  is defined to be  $F_{(X,Y)}(x, y) = \mathbb{P}(X \leq x, Y \leq y)$ .

**Proposition 5.2.** For a fixed  $x \in \mathbb{R}$ , the following holds.

- $\lim_{y \rightarrow \infty} F_{(X,Y)}(x, y) = F_X(x)$
- $\lim_{y \rightarrow -\infty} F_{(X,Y)}(x, y) = 0$
- $\lim_{x, y \rightarrow \infty} F_{(X,Y)}(x, y) = 1$
- $\lim_{\varepsilon \rightarrow 0^+} F_{(X,Y)}(x, y + \varepsilon) = F_{(X,Y)}(x, y)$

**Definition 5.3** (Joint and marginal PMF). Let  $X \in \{x_i \mid i \in \mathbb{N}\}$  and  $Y \in \{y_j \mid j \in \mathbb{N}\}$  be discrete RVs, the joint PMF is defined to be  $f_{(X,Y)}(x_i, y_j) = \mathbb{P}(X = x_i, Y = y_j)$ . We can retrieve the marginal PMF of  $X$  by fixing  $x$  and summing over the possible values of  $Y$ .

$$f_X(x_i) = \sum_{j=0}^{\infty} f_{(X,Y)}(x_i, y_j)$$

**Definition 5.4** (Joint and marginal PDF). We say that  $(X, Y)$  is a continuous multivariate RV if there exists a nonnegative function  $f_{(X,Y)}$  such that for any  $(x, y) \in \mathbb{R}^2$ , we have

$$F_{(X,Y)}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f_{(X,Y)}(u, v) dv du$$

This function is called the joint PDF and we can retrieve the marginal PDF for  $X$  by fixing some  $x$  and integrating for  $y$  on  $\mathbb{R}$ .

$$f_X(x) = \int_{\mathbb{R}} f_{(X,Y)}(x, y) dy$$

## 5.2 Transformations

**Definition 5.5.** Let  $X$  and  $Y$  be discrete RVs and  $g : \mathbb{R}^2 \rightarrow \mathbb{R}$  be a Borel function, then  $g(X, Y)$  is a RV taking values in  $\{z_k \mid k \in \mathbb{N}\}$ . We define its PMF like so:

$$\mathbb{P}(g(X, Y) = z_k) = \sum_{\{(i,j) \mid g(x_i, y_j) = z_k\}} \mathbb{P}(X = x_i, Y = y_j)$$

**Definition 5.6.** Let  $(X, Y)$  be a continuous RV and  $g : \mathbb{R}^2 \rightarrow \mathbb{R}$  be a Borel function, then  $g(X, Y)$  is also a RV with the following CDF:

$$\mathbb{P}(g(X, Y) \leq z) = \int \int_{\{(x,y) \mid g(x,y) \leq z\}} f_{(X,Y)}(x, y) dy dx$$

**Proposition 5.7.** Let  $(X, Y)$  be a continuous RV and  $f_1, f_2 : \mathbb{R} \rightarrow \mathbb{R}$  be two Borel functions with  $U = f_1(X, Y)$  and  $V = f_2(X, Y)$ . If  $(f_1, f_2)^{-1} = (g_1, g_2)$  is the unique inverse,  $g_1$  and  $g_2$  have continuous partial derivatives and  $\det \frac{\partial(g_1, g_2)}{\partial(u, v)} \neq 0$ , then  $(U, V)$  is a continuous RV and the PDF is defined like so:

$$f_{(U, V)}(u, v) = f_{(X, Y)}(g_1(u, v), g_2(u, v)) \left| \det \frac{\partial(g_1, g_2)}{\partial(u, v)} \right|$$

**Examples 5.8.** Here are two common transformations we have seen multiple times.

- If  $f_1(x, y) = x + y$  and  $f_2(x, y) = x - y$ , then  $f_{(U, V)}(u, v) = f_{(X, Y)}\left(\frac{u+v}{2}, \frac{u-v}{2}\right) \cdot \frac{1}{2}$ .
- If  $f_1(x, y) = \sqrt{x^2 + y^2}$  and  $f_2 = \arctan \frac{y}{x}$ , then  $f_{(U, V)}(u, v) = u f_{(X, Y)}(-u \cos v, -u \sin v)$ .

**Proposition 5.9 (Consequences of independence).** Let  $X$  and  $Y$  be RVs,  $X \perp\!\!\!\perp Y$ , i.e.  $F_{(X, Y)}(x, y) = F_X(x)F_Y(y)$  and  $f, g : \mathbb{R} \rightarrow \mathbb{R}$  be Borel functions, then the following holds.

- $f_{(X, Y)}(x, y) = f_X(x)f_Y(y)$
- $M_{(X, Y)}(s, t) = M_X(s)M_Y(t)$
- $f(X) \perp\!\!\!\perp g(Y)$
- $\mathbb{E}[f(X)g(Y)] = \mathbb{E}[f(X)]\mathbb{E}[g(Y)]$
- If the second moments are defined for both  $X$  and  $Y$ , then  $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$

**Definition 5.10 (i.i.d. sequence).** Let  $\{X_i \mid i \in \mathbb{N}\}$  be a sequence of RVs, we say that they are independent and identically distributed if the sequence is an independent family and  $\forall i \in \mathbb{N}, F_{X_i} = F_{X_1}$ .

**Examples 5.11.** We calculated the sum of notable independent sequences of RVs.

- Let  $X_i$  be independent  $B(n_i, p)$  RVs, then  $\sum_{i=0}^m X_i$  is a  $B(\sum_{i=0}^m n_i, p)$  RV.
- Let  $X_i$  be independent  $P(\lambda_i)$  RVs, then  $\sum_{i=0}^m X_i$  is a  $P(\sum_{i=0}^m \lambda_i)$  RV.
- Let  $X_i$  be i.i.d.  $\exp(\lambda)$  RVs, then  $\sum_{i=0}^m X_i$  is a  $\Gamma(m, \lambda)$  RV.
- Let  $X_i$  be i.i.d.  $N(m, \sigma^2)$  RVs, then  $\sum_{i=0}^m X_i$  is a  $N(nm, n\sigma^2)$  RV.

### 5.3 Covariance and Correlation

Let  $X$  and  $Y$  be RVs, we give the following definitions and formal consequences.

**Definition 5.12** (Covariance).

$$\text{Cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

**Proposition 5.13.**

$$|\text{Cov}(X, Y)| \leq \sqrt{\text{Var}(X) \text{Var}(Y)}$$

**Definition 5.14** (Correlation).

$$\text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X) \text{Var}(Y)}}$$

**Proposition 5.15.**

$$\text{Corr}(X, Y) = 1 \Leftrightarrow \exists c > 0, Y = cX$$

**Proposition 5.16.**

$$X \perp\!\!\!\perp Y \implies \text{Cov}(X, Y) = 0 = \text{Corr}(X, Y)$$

**Proposition 5.17.** Let  $\{X_i \mid i \in \mathbb{N}\}$  be an uncorrelated sequence of variables such that  $\mathbb{E}[X_i] = m$  and  $\text{Var}(X_i) = \sigma^2$  for any  $i$ , with  $m \in \mathbb{R}$  and  $\sigma \in (0, \infty)$ . Let  $S_n = \sum_{i=0}^n X_i$  and  $\overline{S}_n = \frac{S_n}{n}$ , the following holds.

- For any  $n \in \mathbb{N}$ ,  $\mathbb{E}[\overline{S}_n] = m$  and  $\text{Var}(\overline{S}_n) = \frac{\sigma^2}{n}$ .
- For any  $\varepsilon > 0$  and  $n \in \mathbb{N}$ ,  $\mathbb{P}(|\overline{S}_n - m| > \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2}$ .

## 6 Conditional Distribution

**Definition 6.1** (Conditional CDF). Let  $X$  and  $Y$  be RVs, the conditional CDF of  $X$  given  $Y = y$  is defined like so:

$$F_{X|Y}(x \mid y) = \lim_{\varepsilon \rightarrow 0} \mathbb{P}(X \leq x \mid Y \in (y - \varepsilon, y + \varepsilon)) = \lim_{\varepsilon \rightarrow 0} \frac{\mathbb{P}(X \leq x, Y \in (y - \varepsilon, y + \varepsilon))}{\mathbb{P}(Y \in (y - \varepsilon, y + \varepsilon))}$$

*Remark 6.2.* Let  $X \in \{x_i \mid i \in \mathbb{N}\}$  and  $Y \in \{y_j \mid j \in \mathbb{N}\}$  be discrete RVs,  $F_{X|Y}(x \mid y)$  is only defined if  $y \in \{y_j \mid j \in \mathbb{N}\}$ . Also, the conditional PMF is defined by  $f_{X|Y}(x_i \mid y_j) = \mathbb{P}(X = x_i \mid Y = y_j)$ .

**Definition 6.3** (Discrete conditional expectation). If  $\sum_{i=0}^{\infty} |x_i| f_{X|Y}(x_i \mid y) < \infty$  for any  $y \in \{y_j \mid j \in \mathbb{N}\}$ , then we can define the conditional expectation which is a RV:

$$\mathbb{E}[X \mid Y](\omega) = \sum_{i=0}^{\infty} |x_i| f_{X|Y}(x_i \mid Y(\omega))$$

**Proposition 6.4.**

$$\mathbb{E}[\mathbb{E}[X | Y]] = \mathbb{E}[X]$$

**Proposition 6.5.** If  $X \perp\!\!\!\perp Y$ , then  $\forall \omega \in \Omega, \mathbb{E}[X | Y](\omega) = \mathbb{E}[X]$ .

**Proposition 6.6.** Let  $(X, Y)$  be a continuous RV with joint PDF  $f_{(X,Y)}$  and marginal PDFs  $f_X$  and  $f_Y$ , then we have the following formulas:

$$F_{X|Y}(x | y) = \int_{-\infty}^x \frac{f_{(X,Y)}(u, y)}{f_Y(y)} du$$

$$f_{X|Y}(x | y) = \frac{f_{(X,Y)}(x, y)}{f_Y(y)}$$

**Definition 6.7** (Continuous conditional expectation). If  $\int_{\mathbb{R}} |x| f_{X|Y}(x | y) dx < \infty$  for any  $y \in \mathbb{R}$ , then  $\mathbb{E}[X | Y](\omega) = \int_{\mathbb{R}} x f_{X|Y}(x | Y(\omega)) dx$ .

**Definition 6.8** (Conditional variance). Let  $(X, Y)$  be a RV with  $\mathbb{E}[X^2] < \infty$  and assume that for every  $\omega \in \Omega$ ,  $\mathbb{E}[X^2 | Y](\omega)$  and  $\mathbb{E}[X | Y](\omega)$  are defined. The conditional variance is defined by  $\text{Var}(X | Y)(\omega) = \mathbb{E}[X^2 | Y](\omega) - (\mathbb{E}[X | Y](\omega))^2$ .

**Proposition 6.9.** In the same setting as last definition, we have the following.

$$\text{Var}(X) = \text{Var}(\mathbb{E}[X | Y]) + \mathbb{E}[\text{Var}(X | Y)]$$

## 7 Asymptotics

Throughout this section, we will consider a sequence of RVs  $\{X_i | i \in \mathbb{N}\}$  with DFs  $\{F_i | i \in \mathbb{N}\}$  and a RV  $X$  with DF  $F$ .

**Definition 7.1** (Convergence in distribution). If for all continuous points  $x$  in  $F$  we have  $\lim_{n \rightarrow \infty} F_n(x) = F(x)$ , then we say that  $X_n \rightarrow X$  in distribution.

**Theorem 7.2.** If  $\forall n, X_n, X \in \mathbb{N}$  then,  $X_n \rightarrow X$  in distribution if and only if for all  $k \in \mathbb{N}$ , we have  $\lim_{n \rightarrow \infty} \mathbb{P}(X_n = k) = \mathbb{P}(X = k)$ .

**Theorem 7.3.** If  $\forall n, X_n$  is continuous with PDF  $f_n$  and  $X$  is continuous with PDF  $f$  and for almost every  $x \in \mathbb{R}$ , we have  $\lim_{n \rightarrow \infty} f_n(x) = f(x)$ , then,  $X_n \rightarrow X$  in distribution.

**Proposition 7.4.** If  $X_n \rightarrow X$  in distribution and  $a, b \in \mathbb{R}$ , then  $aX_n + b \rightarrow aX + b$  in distribution.

**Definition 7.5** (Convergence in probability). If for every  $\varepsilon > 0$ , we have  $\lim_{n \rightarrow \infty} \mathbb{P}(|X_n - X| > \varepsilon) = 0$ , then we say that  $X_n \rightarrow X$  in probability.

**Proposition 7.6.** If  $X_n \rightarrow X$ ,  $X_n \rightarrow X'$  and  $Y_n \rightarrow Y$  in probability then the following holds.

- $X_n \rightarrow X$  in distribution.

- $\mathbb{P}(X = X') = 1$ .
- $X_n + Y_n \rightarrow X + Y$  and  $X_n - Y_n \rightarrow X - Y$  and  $X_n Y_n \rightarrow XY$  in probability.

**Theorem 7.7** (Slutsky). *If  $X_n \rightarrow X$  in distribution and  $Z_n \rightarrow c$  in probability (where  $c \in \mathbb{R}$  is viewed as a Dirac RV at  $c$ ), then  $X_n + Z_n \rightarrow X + c$  and  $X_n Z_n \rightarrow cX$  in distribution.*

**Proposition 7.8.** *If  $X_n \rightarrow c$  in distribution and  $g : \mathbb{R} \rightarrow \mathbb{R}$  is a Borel function, the following holds.*

- $X_n \rightarrow c$  in probability.
- $g(X_n) \rightarrow g(c)$  in distribution.
- $\lim_{n \rightarrow \infty} \mathbb{E}[g(X_n)] = g(c)$

In the following, we denote  $S_n = \sum_{i=0}^n X_i$  and  $\bar{S}_n = \frac{S_n}{n}$  to be the partial sum of the RVs and the rescaled partial sum respectively.

**Definition 7.9.** We say that the sequence  $\{X_i \mid i \in \mathbb{N}\}$  obeys the weak law of large numbers (WLLN) if  $\frac{S_n - \mathbb{E}[S_n]}{n} \rightarrow 0$  in probability. Equivalently,  $\bar{S}_n - \mathbb{E}[\bar{S}_n] \rightarrow 0$  in probability.

**Theorem 7.10** (WLLN-I). *If  $\{X_i\}$  is uncorrelated and the second moments are bounded (i.e.  $\sup_{i \in \mathbb{N}} \mathbb{E}[X_i^2] = M < \infty$ ), then the sequence obeys WLLN.*

**Theorem 7.11** (WLLN-II). *If  $\{X_i\}$  is pairwise independent and identically distributed with  $\mathbb{E}[X_i] = m$  for any  $i$ , then the sequence obeys WLLN.*

**Theorem 7.12** (Central Limit Theorem). *If  $\{X_i\}$  is i.i.d. with  $\mathbb{E}[X_i] = m$  and  $\text{Var}(X_i) = \sigma^2$ , then  $\frac{S_n - mn}{\sqrt{n\sigma^2}} \rightarrow X$  in distribution where  $X$  a  $N(0, 1)$  RV.*