This document contains a summery of the MVE055 exam on the 30th of October 2018, by ____________

Independence
$\Pr(A \cap B) = \Pr(A)\Pr(B)$
Bays Theorem
$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$
$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{j=0} P(B|A_j)P(A_j)}$
Law of total probability
$\Pr(B) = \sum_j \Pr(B|A_j)\Pr(A_j)$
Density function
A function f is a density function of a random variable only if $f(x) \geq 0\ \forall x$ and $\sum_{allx} f(x) = 1$
Expected value
Discrete expected value $\mathbb{E}[X] = \sum_{allx} xf(x)$
$\mu = \mathbb{E}[c] = c$
$\mathbb{E}[cX] = c\mathbb{E}[X]$
$\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$
$\mathrm{Var}[X] = \mathbb{E}[(X-\mu)^2]$
Variance and Standard deviation
$\sigma = \mathrm{Sd}[X] = \sqrt{\mathrm{Var}[X]}$
$\mathrm{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$
$\mathrm{Var}[c] = 0$
$\mathrm{Var}[cX] = c^2\mathrm{Var}[X]$
$\mathrm{Sd}[cX] = c\mathrm{Sd}[X]$
$\mathrm{Var}(X) = \mathrm{Cov}(X,X)$
$\mathrm{Var}(X) = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2$
Bernoulli

$$f(x) = \begin{cases} p, & \text{if } x = 1 \\ 1-p, & \text{if } x = 0 \\ 0, & \text{otherwise} \end{cases}$$

$\mathbb{E}[X] = p$
$\mathrm{Var}[X] = p(1-p)$
Important series
For $|s| < 1$ it holds:
$\sum_{k=0}^{\infty} s^k = \frac{1}{1-s}$
$\sum_{k=0}^{n} s^k = \frac{1-s^{n+1}}{1-s}$
$(a+b)^n = \sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k}$
Geometric random variable

$$f(x) = \begin{cases} p(1-p)^{x-1}, & \text{if } x \in \mathbb{N} \\ 0, & \text{otherwise} \end{cases}$$

$\mathbb{E}[X] = \frac{1}{p}$
$\mathrm{Var}[X] = \frac{1-p}{p^2}$
Binomial random variable

$$f(x) = \begin{cases} \binom{n}{x} p^x(1-p)^{n-x}, & \text{if } x \in \mathbb{N} \\ 0, & \text{otherwise} \end{cases}$$

$\mathbb{E}[X] = np$
$\mathrm{Var}[X] = np(1-p)$
* Constant p, * independent
$X \sim Bin(\#total, P(success))$
Moment Generating Functions
The moments of a r.v. are $\mathbb{E}[X], \mathbb{E}[X^2], \ldots$
$m_X(t) = \mathbb{E}[e^{tX}]$
Provided the RHS is finite $\forall t$ in some open interval.

Normal distribution
X is this if $\mu \in \mathbb{R}$ and $\sigma > 0$ if
$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{\frac{(x-\mu)^2}{2\sigma^2}}, \forall x \in \mathbb{R}$
Notated as $X \sim N(\mu, \sigma^2)$
$\mathbb{E}[X] = \mu$
$\mathrm{Var}[X] = \sigma^2$
Standardization
$X \sim N(\mu,\sigma^2)$ then $\frac{X-\mu}{\sigma} \sim N(0,1)$.
Joint distribution
bivariate discrete random variable
$f_{XY}(x,y) = \Pr[X=x, Y=y], \forall (x,y) \in \mathbb{R}^2$
joint density for the vector (X,Y)
$f(x,y) \geq 0$
$\sum_{all(x,y)} f(x,y) = 1$
The marginal density f
$f_X(x) = \sum_{ally} f_{XY}(x,y)$
$f_Y(x) = \sum_{allx} f_{XY}(x,y)$
Two X and Y with joint f_{XY} and marginal densities f_X,f_Y is independent iff
$f_{XY}(x,y) = f_X(x)f_Y(y), \forall x \forall y$
Expected value
$\mathbb{E}[H(X,Y)] = \sum_{all(x,y)} H(x,y)f_{XY}(x,y)$
Iff X,Y is independent $\Rightarrow$ $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$
Definition: Covariance
$\mathrm{Cov}(X,Y) = \mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])]$
$\mathrm{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$
Iff X,Y is independent $\Rightarrow$
$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y] \Rightarrow \mathrm{Cov}(X,Y) = 0$
$\mathrm{Cov}(X,Y)$ indicates association of X, Y
$\mathrm{Cov}(X,Y) \in \mathbb{R} \Rightarrow$ no information about the strength of the dependence?
Definition (Correlation)
$\rho_{XY} = \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}[X]\mathrm{Var}[Y]}$
ρ_{XY} mesures linear dependance and $\rho_{XY} \in [-1,1]$
$|\rho_{XY}| = 1$ iff $Y = \beta_0 + \beta_1 X \text{where} \beta_0, \beta_1 \neq 0$
Definition (conditional density) X,Y with f_{XY} and f_X, f_Y, then for X with $Y = y$ the
$f_{X|y} = \frac{f_{XY}(x,y)}{f_Y(y)}$, if $f_Y(y) > 0$
Poisson distribution
$f(x) = \frac{e^{-\lambda}\lambda^x}{x!}, x \in \mathbb{N}, \lambda > 0$
$\mathbb{E}[X] = \lambda$, $\mathrm{Var}(X) = \lambda$,
λ is avg events in 1unit of time
Iff $X_1 \sim Poisson(\lambda_1)$, $X_2 \sim Poisson(\lambda_2)$ are independent then $X_1 + X_2 \sim Poisson(\lambda_1 + \lambda_2)$
Hypergeometric distribution
$f(x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}}$
$\max\{0, n-(N-r)\} \leq x \leq \min\{n, r\}$
$\mathbb{E}[X] = n\frac{r}{N}$, $\mathrm{Var}(X) = n\frac{r}{N}\frac{N-r}{N}\frac{N-n}{N-1}$
We select n objects from N objects, of which r has a trait. X counts how many of the selected objects have the trait
Exponential distribution
$f(x) = \frac{1}{\beta}e^{-\frac{x}{\beta}}$
$\mathbb{E}[X] = \beta$
$\mathrm{Var}(X) = \beta^2$
$F_X(X) = 1 - e^{-\frac{x}{\beta}}$, $\lambda = \frac{1}{\beta}$
negative Binomial distribution
$p \in [1,0]$, $r \in \mathbb{N}\neg\{0\}$
$f(x) = \binom{x-1}{r-1}(1-p)^{x-r}p^r, x = r, r+1, r+2, \ldots$
$\mathbb{E}[X] = \frac{r}{p}$, $\mathrm{Var}(X) = \frac{r(1-p)}{p^2}$
Consider a sequence of independent and identical experiments, each one with probability p of success. X models the number of trails needed to obtain r successes.
Generating function
given $\{a_n\}_{n=0}^{\infty}$ then $g(x) = \sum_{n=0}^{\infty} a_n x^n$
let $a_n = c^n$ then $\sum_{n=0}^{\infty} c^n x^n = \sum_{n=0}^{\infty}(cx)^n = \frac{1}{1-cx}$
$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \prod_{i=1}^{k}\frac{n-i+1}{i}$
See binom theom: $\sum_{k=0}^{n}\binom{n}{k}x^k = (1+x)^n$
Let $a_n = \binom{n+k}{k}$ then $\sum_{n=0}^{\infty} a_n x^n = \frac{1}{(1-x)^{k+1}}$
Chebychev's inequality:
Let $X \ni \mathbb{E} = \mu \wedge \mathrm{Var}(X) = \sigma^2$, iff $0 < \sigma^2 < \infty$ then $\forall a > 0$ it holds $P(|X-\mu| \geq a) \leq \frac{\sigma^2}{a^2}$
Let $Y = |X-\mu|$. $\forall a > 0$ define

$$Z = \begin{cases} a^2, & Y \geq a \\ 0, & \text{otherwise} \end{cases}$$

then $Z \leq Y^2 \Rightarrow \mathbb{E}[Z] \leq \mathbb{E}[Y^2](1)$
see $\mathbb{E}[Z] = a^2 P(Y \geq a) = a^2 P(|X-\mu| \geq a)$
see $\mathbb{E}[Y^2] = \mathbb{E}[|X-\mu|^2] = \mathrm{Var}(X) = \sigma^2$
substitute into (1): $a^2 P(|X-\mu| \geq a) \geq \mathrm{Var}(X) \Leftrightarrow P(|X-\mu| \geq a) \geq \frac{\mathrm{Var}(X)}{a^2} = \frac{\sigma^2}{a^2}$
Marcov chain
$P(X_n = x_n | X_{n-1} = x_{n-1}) = P(X_n = x_n | X_{n-1} = x_{n-1}, \ldots, X_0 = x_0)$
$\mathrm{P} = \begin{pmatrix} \mathrm{Q} & \mathrm{R} \\ 0 & \mathrm{I} \end{pmatrix}$

$N = (I-Q)^{-1}$ then $B = NR$
Characteristic function
$\phi_X(t) = \mathbb{E}[e^{itX}]$, $\phi(0) = 1$, $|\phi(t)| \leq 1$
Wilcoxs?
$\mathbb{E}(W_m) = \frac{m(m+n+1)}{2}$
$\mathrm{Var}(W_m) = \frac{m(m+n+1)}{12}$
$\frac{W_m - E(W_m)}{\sqrt{\mathrm{Var}(W_m)}} \sim N(0,1)$
Central limit theorem
Confidence Interval
The $100(1-\alpha)\%$ for $\forall \mu$ is given by
$I_\alpha = [\bar{X} - z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}, \bar{X} + z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}]$
where $z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$ is the point such that $P(N(0,1)) \geq z_{\frac{\alpha}{2}}) = \frac{\alpha}{2}$. The interval I_a has width of $I_a = 2z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$
Iff $\alpha \to 1$, $I_\alpha \to 0$ visavi.
Misc
Given two events A, B (if $\Pr[B] > 0$) we have $\Pr[A|B] = \frac{\Pr[A \cap B]}{\Pr[B]}$,
$X \geq 1$
$\mathbb{E}[U] = \frac{b+a}{2} \forall U \in [a,b]$
then $\mathbb{E}[V] = s + t \cdot \frac{b+a}{2} \forall V = s + tU$
Demorgans law
$P(A^c \cap B^c) = P((A \cup B)^c)$.