FORMULAE SHEET
for
Probability and Statistics
FS 2020

1 Some discrete distributions

1. **Binomial distribution**

   $X \sim \mathrm{Bin}(n, p)$ for $n \in \mathbb{N}$, $p \in [0,1]$:

   $$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k} \qquad \text{for } k \in \{0, 1, \ldots, n\},$$

   $\mathbb{E}[X] = np$, $\mathrm{Var}[X] = np(1-p)$.

2. **Geometric distribution**

   $X \sim \mathrm{Geom}(p)$ for $p \in (0,1)$:

   $$p_X(k) = (1-p)^{k-1} p \qquad \text{for } k \in \{1, 2, \ldots\},$$

   $\mathbb{E}[X] = \frac{1}{p}$, $\mathrm{Var}[X] = \frac{1-p}{p^2}$.

3. **Poisson distribution**

   $X \sim \mathrm{Poi}(\lambda)$ for $\lambda > 0$:

   $$p_X(k) = e^{-\lambda} \frac{\lambda^k}{k!} \qquad \text{for } k \in \{0, 1, 2, \ldots\},$$

   $\mathbb{E}[X] = \lambda$, $\mathrm{Var}[X] = \lambda$.

4. **Hypergeometric distribution**

   $X \sim \mathrm{Hypergeom}(n, N, K)$ for n, N, K positive integers such that $\max(n, K) \leq N$:

   $$p_X(k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$$
   $$\text{for } k \in \{\max(0, n-N+K), \ldots, \min(n, K)\},$$

   $\mathbb{E}[X] = n\frac{K}{N}$, $\mathrm{Var}[X] = n\frac{K}{N}(1 - \frac{K}{N})\frac{N-n}{N-1}$.

2 Some continuous distributions

1. **Uniform distribution**

   $X \sim \mathcal{U}(a, b)$ for $a < b$:

   $$f_X(x) = \frac{1}{b-a} I_{[a,b]}(x),$$

   $\mathbb{E}[X] = \frac{a+b}{2}$, $\mathrm{Var}[X] = \frac{(b-a)^2}{12}$.

2. **Exponential distribution**

   $X \sim \mathrm{Exp}(\alpha)$ for $\alpha > 0$:

   $$f_X(x) = \alpha e^{-\alpha x} I_{(0,\infty)}(x),$$

   $\mathbb{E}[X] = \frac{1}{\alpha}$, $\mathrm{Var}[X] = \frac{1}{\alpha^2}$.

3. **Gamma distribution**

   $X \sim \mathrm{G}(\beta, \alpha)$ for $\beta, \alpha > 0$:

   $$f_X(x) = \frac{\alpha^\beta}{\Gamma(\beta)} x^{\beta-1} e^{-\alpha x} I_{(0,\infty)}(x),$$

   $\mathbb{E}[X] = \frac{\beta}{\alpha}$, $\mathrm{Var}[X] = \frac{\beta}{\alpha^2}$.

4. **Normal distribution**

   $X \sim \mathcal{N}(\mu, \sigma^2)$ for $\mu \in \mathbb{R}$, $\sigma > 0$:

   $$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad \text{for } x \in \mathbb{R},$$

   $\mathbb{E}[X] = \mu$, $\mathrm{Var}[X] = \sigma^2$.

   If $\mu = 0$ and $\sigma = 1$, we say that X is standard normal and its distribution function (cdf) is denoted by Φ.

3 Some inequalities

1. **Markov's inequality**

   Let X be a random variable and $g : \mathbb{R} \to [0, \infty)$ increasing. For any c with $g(c) > 0$,

   $$\mathbb{P}[X \geq c] \leq \frac{\mathbb{E}[g(X)]}{g(c)}.$$

2. **Chebyshev's inequality**

   Let X be a random variable whose expectation exists and with finite variance. For any $c > 0$,

   $$\mathbb{P}\big[|X - \mathbb{E}[X]| \geq c\big] \leq \frac{\mathrm{Var}[X]}{c^2}.$$

3. **Cauchy–Schwarz inequality**

   Let X, Y be random variables defined on the same probability space with $\mathbb{E}[X^2], \mathbb{E}[Y^2]$ both finite. Then

   $$|\mathbb{E}[XY]| \leq \sqrt{\mathbb{E}[X^2]\,\mathbb{E}[Y^2]}.$$

4. **Jensen's inequality**

   Let X be a random variable with $\mathbb{E}[|X|] < \infty$ and $g : \mathbb{R} \to \mathbb{R}$ a convex function. Then

   $$\mathbb{E}[g(X)] \geq g(\mathbb{E}[X]).$$

   If g is strictly convex, then equality holds if and only if $\mathbb{P}[X = \mathbb{E}[X]] = 1$.

4 Miscellaneous

1. **Covariance**

   For two random variables X, Y defined on the same probability space with $\mathbb{E}[|X|], \mathbb{E}[|Y|], \mathbb{E}[|XY|]$ all finite,

   $$\begin{aligned} \mathrm{Cov}(X,Y) &= E\big[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])\big] \\ &= \mathbb{E}[XY]-\mathbb{E}[X]\mathbb{E}[Y]. \end{aligned}$$

   Cov is bilinear and symmetric.

2. **Correlation**

   For two random variables X, Y defined on the same probability space with $\mathrm{Var}[X], \mathrm{Var}[Y]$ both finite,

   $$\mathrm{corr}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}[X]\mathrm{Var}[Y]}}.$$

   We always have $-1 \leq \mathrm{corr}(X,Y) \leq 1$, with equality if and only if there exist $a \neq 0$ and $b \in \mathbb{R}$ with $\mathbb{P}[Y = aX + b] = 1$ and $\mathrm{sign}(a) = \mathrm{sign}(\mathrm{corr}(X,Y))$.

5 Some test statistics

Suppose $X_1, \ldots, X_n$ are i.i.d. random variables with an appropriate distribution (depending on the desired test). Suppose $Y_1, \ldots, Y_m$ are i.i.d. random variables which are independent of $X_1, \ldots, X_n$ and with an appropriate distribution (depending on the desired test). Set

$$\begin{aligned} \overline{X}_n &= \frac{1}{n}\sum_{i=1}^n X_i, \qquad & S_X^2 &= \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X}_n)^2, \\ \overline{Y}_m &= \frac{1}{m}\sum_{k=1}^m Y_k, \qquad & S_Y^2 &= \frac{1}{m-1}\sum_{k=1}^m (Y_k - \overline{Y}_m)^2. \end{aligned}$$

1. **One-sample t-test**

   $$T = \frac{\overline{X}_n - \mu_0}{S_X/\sqrt{n}}.$$

2. **z-test**

   $$T = \frac{\overline{X}_n - \mu_0}{\sigma/\sqrt{n}}.$$

3. **Two-sample t-test**

   $$T = \frac{\overline{X}_n - \overline{Y}_m}{\sqrt{\frac{1}{n}+\frac{1}{m}}\sqrt{\frac{1}{n+m-2}((n-1)S_X^2 + (m-1)S_Y^2)}}.$$

4. **Sign test**

   $$T = \sum_{i=1}^n I_{\{X_i > \mu_0\}}.$$

5. **Two-sample Wilcoxon test**

   $$T = \sum_{i=1}^n \sum_{k=1}^m I_{\{X_i < Y_k\}}.$$

6. **Goodness-of-fit χ^2-test**

   $$T = \sum_{i=1}^k \frac{(N_i - n\theta_{0,i})^2}{n\theta_{0,i}}.$$

In all cases, the distribution of the test statistic T under the null hypothesis is known (exactly or approximately), and the relevant quantiles can be found in the corresponding tables.