STATISTICS 743 (Part I)

The following definitions of common probability distributions, their means and variances may be used without proof in your solutions.

The probability mass or density functions are 0 outside of the values or ranges of values specified below.

- The binomial(n, p) probability mass function is

$$f(x \mid p) = \binom{n}{x} p^x (1-p)^{n-x} \qquad x = 0, 1, \ldots, n$$

  where n is a positive integer and $0 < p < 1$.

  If $X \sim$ binomial(n, p) then $\mathrm{E}(X) = np$ and $\mathrm{Var}(X) = np(1-p)$.

- The Poisson(λ) probability mass function is

$$f(x \mid \lambda) = \frac{\lambda^x \mathrm{e}^{-\lambda}}{x!} \qquad x = 0, 1, \ldots$$

  where $\lambda > 0$

  If $X \sim$ Poisson(λ) then $\mathrm{E}(X) = \lambda$ and $\mathrm{Var}(X) = \lambda$.

- The Geometric(p) probability mass function is

$$f(x \mid p) = p(1-p)^{x-1} \qquad x = 1, 2, \ldots$$

  where $0 < p < 1$.

  If $X \sim$ geometric(p) then $\mathrm{E}(X) = 1/p$ and $\mathrm{Var}(X) = (1-p)/p^2$.

- The Negative Binomial(r, p) probability mass function is

$$f(x \mid r, p) = \binom{r+x-1}{x} p^r (1-p)^x \qquad x = 0, 1, \ldots$$

  where r is a positive integer and $0 < p < 1$.

  If $X \sim$ negative binomial(r, p) then $\mathrm{E}(X) = r(1-p)/p$ and $\mathrm{Var}(X) = r(1-p)/p^2$.

- The uniform(a, b) probability density function is

$$f(x \mid a, b) = \frac{1}{b-a} \qquad a < x < b.$$

where a and b are real numbers with $b > a$.

If $X \sim$ uniform(a, b) then $\mathrm{E}(X) = (a+b)/2$ and $\mathrm{Var}(X) = (b-a)^2/12$.

- The univariate normal(μ, σ^2) probability density function is

$$f(x \mid \mu, \sigma^2) \;=\; \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} \qquad x \in \mathbb{R}$$

where $\mu \in \mathbb{R}$ and $\sigma^2 > 0$.

If $X \sim$ normal(μ, σ^2) then $\mathrm{E}(X) = \mu$ and $\mathrm{Var}(X) = \sigma^2$.

- The exponential(μ) probability density function is

$$f(x; \mu) = \frac{1}{\mu} \mathrm{e}^{-x/\mu} \qquad x \geqslant 0.$$

where $\mu > 0$.

If $X \sim$ exponential(μ) then $\mathrm{E}(X) = \mu$ and $\mathrm{Var}(X) = \mu^2$.

- The gamma(α, β) probability density function is

$$f(x; \alpha, \beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} \mathrm{e}^{-x/\beta} \qquad x \geqslant 0.$$

where $\alpha > 0$ and $\beta > 0$ and the gamma function is defined as

$$\Gamma(\alpha) \;=\; \int_0^\infty x^{\alpha-1} \mathrm{e}^{-x}\, dx$$

If $X \sim$ gamma(α, β) then $\mathrm{E}(X) = \alpha\beta$ and $\mathrm{Var}(X) = \alpha\beta^2$.

- If a random variable X has a gamma distribution with $\beta = 2$ and $\alpha = p/2$ then we say that X has a chi-squared distribution with p degrees of freedom.

- The beta(α, β) probability density function is

$$f(x; \alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} \qquad 0 < x < 1.$$

where $\alpha > 0$ and $\beta > 0$.

If $X \sim$ beta(α, β) then $\mathrm{E}(X) = \alpha/(\alpha+\beta)$ and $\mathrm{Var}(X) = \alpha\beta/\big((\alpha+\beta)^2(\alpha+\beta+1)\big)$.

The following results may be used without proof in your solutions

- If A and B are any two events then $\Pr(A \bigcup B) \;=\; \Pr(A) + \Pr(B) - \Pr(A \bigcap B)$.

- If $A_1, A_2, \ldots$ is a partition of a sample space and B is any event in the same sample space then
$$P(A_i \mid B) = \frac{P(B \mid A_i)P(A_i)}{\sum_{j=1}^{\infty} P(B \mid A_j)P(A_j)}$$

- If X is a random variable then the moment generating function of X is defined to be $M_X(t) = \mathrm{E}\left(\mathrm{e}^{tX}\right)$ provided this expectation exists for t in a neighbourhood of 0.

- $\lim_{n\to\infty} \left(1 + \frac{a_n}{n}\right)^n \;=\; \exp\left\{\lim_{n\to\infty} a_n\right\}$.

- A distribution is said to be in the exponential family if its probability density or mass function can be written as
$$f(x \mid \boldsymbol{\theta}) = h(x)c(\boldsymbol{\theta}) \exp\left\{\sum_{i=1}^{k} w_i(\boldsymbol{\theta})t_i(x)\right\}$$
for positive function $h(x)$ and $c(\boldsymbol{\theta})$.

- A distribution is said to be in a location-scale family if its probability density function can be written as
$$f(\mid \mu, \sigma) \;=\; \frac{1}{\sigma} g\left(\frac{x-\mu}{\sigma}\right)$$
for some density function g.

- If X and Y are two random variables with joint density (or mass) function $f_{X,Y}$ then the conditional density (or mass) function is defined to be
$$f_{Y|X}(y \mid X = x) \;=\; \frac{f_{X,Y}(x,y)}{f_X(x)}$$
where f_X is the marginal pdf or pmf for X obtained by integrating or summing $f_{X,Y}$ over all possible values of y.

- For any two random variables X and Y we can write $\mathrm{E}(X) \;=\; \mathrm{E}\big(\mathrm{E}(X \mid Y)\big)$ and $\mathrm{Var}(X) \;=\; \mathrm{E}\big(\mathrm{Var}(X \mid Y)\big) + \mathrm{Var}\big(\mathrm{E}(X \mid Y)\big)$.

- The correlation coefficient between two random variables X and Y is
$$\rho_{X,Y} \;=\; \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}$$
where the covariance is $\mathrm{Cov}(X,Y) = \mathrm{E}\big((X - \mathrm{E}(X))(Y - \mathrm{E}(Y))\big) = \mathrm{E}(XY) - \mathrm{E}(X)\mathrm{E}(Y)$.

- The bivariate normal distribution has joint pdf

$$\begin{aligned} f(x_1, x_2) &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \\ &\quad \times \exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{(x_1-\mu_1)^2}{\sigma_1^2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2} - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2}\right)\right\} \end{aligned}$$

  where μ_1 and μ_2 are the marginal means, σ_1^2 and σ_2^2 are the marginal variances and $-1 < \rho < 1$ is the correlation between the two components of the random vector.

- If $X_1, \ldots, X_n$ are random variables and $a_1, \ldots, a_n$ are constants then

$$\begin{aligned} \mathrm{E}\left(\sum_{i=1}^{n} a_i X_i\right) &= \sum_{i=1}^{n} a_i \mathrm{E}(X_i) \\ \mathrm{Var}\left(\sum_{i=1}^{n} a_i X_i\right) &= \sum_{i=1}^{n} a_i^2 \mathrm{Var}(X_i) + 2\sum_{i>j} a_i a_j \mathrm{Cov}(X_i, X_j) \end{aligned}$$

- $X_n \xrightarrow{d} X$ if $\lim_{n\to\infty} \Pr(X_n \leqslant x) = \Pr(X \leqslant x)$ at every point x at which the cumulative distribution function of X is continuous.

- $X_n \xrightarrow{p} X$ if, for every $\varepsilon > 0$, $\lim_{n\to\infty} \Pr(|X_n - X| < \varepsilon) = 1$.

- $X_n \xrightarrow{p} X \;\Rightarrow\; X_n \xrightarrow{d} X$.

- A sequence $X_1, X_2, \ldots,$ is bounded in probability if for every $\varepsilon > 0$ there exist constants B_ε and N_ε such that $n > N_\varepsilon \;\Rightarrow\; \Pr(|X_n| < B_\varepsilon) \geqslant 1 - \varepsilon$.

- If $U \sim \mathrm{Uniform}(0, 1)$ and F is a cumulative distribution function with inverse F^{-1} then $X = F^{-1}(U)$ has cumulative distribution function F.

- A statistic $T(\boldsymbol{X})$ is sufficient for the parameter θ if the joint pdf (or pmf) of $\boldsymbol{X}$ can be written as $f(\boldsymbol{x} \mid \theta) = g(T(\boldsymbol{x}), \theta)h(\boldsymbol{x})$ for every $\boldsymbol{x}$.

- The family of sampling distributions of a statistic T is complete if
  $\mathrm{E}_\theta(g(T)) = 0 \Rightarrow \Pr_\theta(g(T) = 0) = 1$ for every possible θ