MATH10282 Introduction to Statistics

Formulas to remember for the final exam in May/June 2019

Given a data set, know how to compute the sample mean, sample variance, sample p quartile, sample first quartile, sample third quartile, sample median and sample range.

$\widehat{\theta}$ **is an unbiased estimator of** θ **if** $E\left(\widehat{\theta}\right) = \theta$.

$\widehat{\theta}$ **is an asymptotically unbiased estimator of** θ **if** $\lim_{n\to\infty} E\left(\widehat{\theta}\right) = \theta$.

The bias of $\widehat{\theta}$ **is** $E\left(\widehat{\theta}\right) - \theta$.

The mean squared error of $\widehat{\theta}$ **is** $E\left[\left(\widehat{\theta} - \theta\right)^2\right]$.

$\widehat{\theta}$ **is a consistent estimator of** θ **if** $\lim\limits_{n\to\infty} E\left[\left(\widehat{\theta} - \theta\right)^2\right] = 0$.

If $X \sim$ **Exp** (a) **then** $E(X) = 1/a$, $Var(X) = 1/a^2$, $f_X(x) = a\exp(-ax)$ **and** $F_X(x) = 1 - \exp(-ax)$.

If $X \sim$ **Uniform** $[a, b]$ **then** $E(X) = (a+b)/2$, $Var(X) = (b-a)^2/12$ **and** $F_X(x) = \frac{x-a}{b-a}$.

When testing $H_0 : \mu = \mu_0$ **versus** $H_1 : \mu \neq \mu_0$ **the Type I error occurs if** H_0 **is rejected when in fact** $\mu = \mu_0$.

When testing $H_0 : \mu = \mu_0$ **versus** $H_1 : \mu \neq \mu_0$ **the Type II error occurs if** H_0 **is accepted when in fact** $\mu \neq \mu_0$.

The significance level is the probability of type I error.

Suppose $X_1, X_2, \ldots, X_n$ **is a random sample from** $N\left(\mu, \sigma^2\right)$, **where** σ **is known. The rejection region for the following tests are**

(i) reject $H_0 : \mu = \mu_0$ **versus** $H_1 : \mu \neq \mu_0$ **if** $\sqrt{n}\left|\overline{X} - \mu_0\right|/\sigma > z_{1-\frac{\alpha}{2}}$;

(ii) reject $H_0 : \mu = \mu_0$ **versus** $H_1 : \mu < \mu_0$ **if** $\sqrt{n}\left(\overline{X} - \mu_0\right)/\sigma < -z_{1-\alpha}$.

Let $\mathbf{X} = (X_1, \ldots, X_n)$, **with** $X_1, \ldots, X_n$ **an independent random sample from a distribution** F_X **with unknown parameter** θ. **Let** $I(\mathbf{X}) = [a(\mathbf{X}), b(\mathbf{X})]$ **denote an interval estimator for** θ.

(i) $I(\mathbf{X})$ **is a** $100(1-\alpha)\%$ **confidence interval if**

$$\Pr\left(a(\mathbf{X}) < \theta < b(\mathbf{X})\right) = 1 - \alpha;$$

(ii) the coverage probability of $I(\mathbf{X})$ **is**

$$\Pr\left(a(\mathbf{X}) < \theta < b(\mathbf{X})\right);$$

(iii) the coverage length of $I(\mathbf{X})$ **is** $b(\mathbf{X}) - a(\mathbf{X})$.

If $X_1, X_2, \ldots, X_n$ **is a random sample from** $N\left(\mu, \sigma^2\right)$ **then** $\overline{X} \sim N\left(\mu, \sigma^2/n\right)$.

If $X_1, X_2, \ldots, X_n$ **is a random sample from** $N\left(\mu, \sigma^2\right)$ **and** S^2 **denotes the sample variance then** $\sqrt{n}\left(\overline{X} - \mu\right)/S \sim t_{n-1}$.

The gamma function is defined by $\Gamma(a) = \displaystyle\int_0^\infty t^{a-1}\exp(-t)dt$.

$\Gamma(n+1) = n!$ **if** n **is a positive integer.**

$\Gamma(a+1) = a\Gamma(a)$ **and** $\Gamma(1/2) = \sqrt{\pi}$.