STAT 105 Exam II
Reference Sheet

Factorial Analysis (Two Factors)

Assuming

- Factor A with levels $1, 2, ..., I$,
- Factor B with levels $1, 2, ..., J$,
- n is the total number of observations,
- n_{ij} is the total number of observations with Factor A at level i and Factor B at level j,
- $n_{i\cdot}$ is the total number of observations with Factor A at level i,
- $n_{\cdot j}$ is the total number of observations with Factor B at level j.
- y_{ijk} is the kth observation where Factor A is at level i and Factor B is at level j.

$y_{\cdot\cdot} = \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} y_{ijk}$ $\quad$ $\bar{y}_{\cdot\cdot} = \frac{1}{n} \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} y_{ijk}$

$\bar{y}_{i\cdot} = \frac{1}{n_{i\cdot}} \sum_{j=1}^{J} \sum_{k=1}^{K} y_{ijk}$ $\quad$ $\bar{y}_{\cdot j} = \frac{1}{n_{\cdot j}} \sum_{i=1}^{I} \sum_{k=1}^{K} y_{ijk}$

Main effect of Factor A at level i	$a_i = \bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot}$
Main effect of Factor B at level j	$b_j = \bar{y}_{\cdot j} - \bar{y}_{\cdot\cdot}$
Fitted Value	$\hat{y}_{ij} = a_i + b_j + \bar{y}_{\cdot\cdot}$

Discrete Random Variables

Probability function	$P[X = x] = f_X(x)$
Cumulative probability function	$P[X \leq x] = F_X(x)$
Expected Value	$\mu = E(X) = \sum_x x f_X(x)$
Variance	$\sigma^2 = Var(X) = \sum_x (x - \mu)^2 f_X(x)$
Standard Deviation	$\sigma = \sqrt{Var(X)}$

Geometric Random Variables

X is the trial count upon which the first successful outcome is observed performing independent trials with probability of success p.

Possible Values	$x = 1, 2, 3, \ldots$
Probability function	$P[X = x] = f_X(x) = p^x(1-p)^{x-1}$
Expected Value	$\mu = E(X) = \frac{1}{p}$
Variance	$\sigma^2 = Var(X) = \frac{1-p}{p^2}$

Joint Distributions and Related Distributions

Joint Probability Function	$P[X = x, Y = y] = f(x, y)$
Marginal Probability Function	$P[X = x] = f_X(x) = \sum_{\text{all } y} f(x, y)$ $P[Y = y] = f_Y(y) = \sum_{\text{all } x} f(x, y)$
Conditional Probability Function	$P[X = x \mid Y = y] = \frac{f(x,y)}{f_Y(y)}$ $P[Y = y \mid X = x] = \frac{f(x,y)}{f_X(x)}$

Binomial Random Variables

X is the number of successful outcomes observed in n independent trials with probability of success p.

Possible Values	$x = 0, 1, 2, \ldots, n$
Probability function	$P[X = x] = f_X(x) = \frac{n!}{(n-x)!x!} p^x (1-p)^{n-x}$
Expected Value	$\mu = E(X) = np$
Variance	$\sigma^2 = Var(X) = np(1-p)$

Continuous Random Variables

Probability density function	$P[a \leq X \leq b] = \int_a^b f_X(x)dx$
Cumulative probability function	$P[X \leq x] = F_X(x) = \int_{-\infty}^{x} f_X(t)dt$
Expected Value	$\mu = E(X) = \int_{-\infty}^{\infty} x f_X(x)dx$
Variance	$\sigma^2 = Var(X) = \int_{-\infty}^{\infty} (x - \mu)^2 f_X(x)dx$
Standard Deviation	$\sigma = \sqrt{Var(X)}$

Normal Random Variables

Let X be a normal random variable with mean μ and variance σ^2.

Probability density function	$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$
Expected Value	$E(X) = \mu$
Variance	$Var(X) = \sigma^2$

Standard Normal Random Variables (Z)

A normal random variable with mean 0 and variance σ^2.
If X is normal(μ,σ^2) then $P[a \leq X \leq b] = P\left[\frac{a-\mu}{\sigma} \leq Z \leq \frac{b-\mu}{\sigma}\right]$

Probability density function	$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$